

## Chapter 8

### Assessment

#### *Chapter at a Glance*

##### **Purposes of Assessment**

##### **Assessment Cycles**

- Short-Cycle Formative Assessment
- Medium-Cycle Assessment
  - End-of-Unit Assessments
  - Interim or Benchmark Assessments
  - Assessing ELD Progress Over Time
- Long-Cycle Assessment
  - Annual/End-of-Year Assessments
  - Ensuring Accessibility for ELs on Long-Cycle Assessments
- Additional Methods of Medium- and Long-Cycle Assessment
  - Rubrics
  - Portfolios

##### **Student Involvement**

- Feedback
  - Teacher Feedback
  - Peer Feedback
  - Self-Assessment

##### **Assessment of ELD Progress**

- Assessing ELD Progress in Writing
- Assessing ELD Progress in Oral Language

##### **Assessment for Intervention**

- Universal Screening (Medium Cycle)
- Diagnostic Assessment (Medium Cycle)
- Progress Monitoring (Short or Medium Cycle)

##### **Mandated California Assessments**

- Computer Adaptive Tests
- Performance Tasks
- Accommodations for Students with Significant Cognitive Disabilities
- English Language Proficiency Assessment

##### **Technical Quality of Assessments**

- Elements of Technical Quality
  - Validity
  - Reliability
  - Freedom from Bias
- Technical Quality and Formative Assessment

##### **Conclusion**

##### **Works Cited**

Student achievement of the California Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects

(CA CCSS for ELA/Literacy) and the California English Language Development Standards (CA ELD Standards) depends on educators' skilled use of assessment information. With these standards, the landscape of assessment and accountability in California has experienced a dramatic shift. Not only do the standards present new goals for California educators as depicted in the outer ring of Figure 8.1 below, but the implementation of the California Assessment of Student Performance and Progress (CAASPP) system represents a major shift in the intent of statewide assessment: "It is the intent of the Legislature...to provide a system of assessments of pupils that has the primary purposes of assisting teachers, administrators, and pupils and their parents; improving teaching and learning; and promoting high-quality teaching and learning using a variety of assessment approaches and types" (E60602.5(a)). This shift is consonant with major emphases in California's standards for college and career readiness: a renewed focus on purposeful and deeper learning for students and their teachers, strong collaboration and partnerships at all levels of education, and a culture of continuous growth based on reflective practice.

Figure 8.1. Goals, Context, and Themes of the CA CCSS for ELA/Literacy and the CA ELD Standards



This chapter describes what is involved in the skilled use of assessment to support student achievement of the CA CCSS for ELA/Literacy and the CA ELD Standards—and ultimately the overarching goals of ELA/literacy and ELD instruction: students develop the readiness for college, careers, and civic life; attain the capacities of literate individuals; become broadly literate; and acquire the skills necessary for living and learning in the 21st century. (See outer ring of Figure 8.1.) Both sets of standards, as discussed throughout the framework, constitute shifts that have significant implications for assessment. First, the CA CCSS for ELA/Literacy and CA ELD Standards are organized in a coherent structure from kindergarten through grade twelve. The CA CCSS for ELA/Literacy within each strand (reading, writing, speaking and listening<sup>1</sup>, and language) can be backward mapped from the CCR Anchor Standards, meaning that students work on a relatively small number of broad competencies to move from novice to expert. Similarly, the CA ELD Standards (Interacting in Meaningful Ways, Learning About How English Works, Using Foundational Literacy Skills) are organized coherently to help teachers make important instruction and assessment decisions for ELs by grade level *and* English language proficiency level. Second, both sets of standards encompass the full spectrum of language and literacy competencies from kindergarten through grade twelve, meaning that students apply and transfer skills from the earliest grades. Third, the standards encourage educators to think broadly and plan instruction comprehensively. “[E]ach standard need not be a separate focus for instruction and assessment. Often, several standards can be addressed by a single, rich task [, so that] students can develop mutually reinforcing skills and exhibit mastery...across a range of texts [and tasks]” (CDE 2013, 4-5). And finally, the standards recommend that language and literacy learning be connected with the academic disciplines from the earliest grades onward. Assessment, then, should enable educators to determine a student’s trajectory in

---

<sup>1</sup> As noted throughout this framework, speaking and listening should be broadly interpreted. Speaking and listening should include deaf and hard of hearing students using American Sign Language (ASL) as they primary language. Students who are deaf and hard of hearing who do not use ASL as their primary language but use amplification, residual hearing, listening and spoken language, cued speech and sign supported speech, access general education curriculum with varying modes of communication.

developing proficiency in language and literacy within and across the years in the standards.

The chapter begins with a discussion of the different purposes of assessment—both *for* and *of* learning. Cycles of assessment—short, medium, and long—are then discussed, including the types and purposes of assessment within each time frame and the decisions that each assessment type can inform. Snapshots of teacher use of assessment are included throughout the discussion of the cycles. The role of student involvement and feedback in assessment is described, followed by guidance for assessment of ELD progress and descriptions of assessment for intervention. In addition, the chapter briefly discusses mandated statewide assessments and concludes with a consideration of the technical quality of assessments to ensure that assessments yield accurate information for their intended purposes.

This chapter can be used in several ways. As a source of professional learning for teachers and school and district leaders, the chapter can play a critical role in strengthening educators' assessment literacy, in other words their knowledge and understanding of assessment practices and appropriate use of assessment evidence to shape powerful instruction. Teachers and leaders can use this chapter to examine the types of assessment practices and sources of assessment evidence currently in use in schools and propose needed additions and adjustments. This chapter highlights the process of formative assessment, which should become the focus of in-depth professional learning and support, including dialogue with peers, time for practicing new approaches, and coaching for reflection and refinement.

### **Purposes of Assessment**

Assessment is designed and used for different purposes. For example, an annual assessment designed to assess how well students have met a specific standard (for example, CA CCSS for ELA/Literacy RI.4.8: *Explain how an author uses reasons and evidence to support particular points in a text*) does just that: It tells educators whether students have met a specific standard. However, it cannot serve the purpose of diagnosing a particular reading difficulty a fourth grade student is experiencing in achieving the standard. Nor can it provide substantive insights into how a student is beginning to understand what constitutes evidence in a specific text. In the use of any

assessment, a central question is, “Am I using this assessment for the purpose for which it is intended?”

Assessment has two fundamental purposes: One is to provide information about student learning minute-by-minute, day-to-day, and week-to-week so teachers can continuously adapt instruction to meet students’ specific needs and secure progress. This type of assessment is intended to assist learning and is often referred to as formative assessment or assessment **for** learning. Formative assessment occurs in real time, during instruction while student learning is underway (Allal 2010; Black and Wiliam 1998; Bell and Cowie 2000; Heritage 2010; Shepard 2000, 2005). For example, a third grade teacher working with small groups of students on distinguishing their point of view from a particular author’s is able to gain insights into students’ developing skills through the use of strategic questions and can adjust instruction and students’ next steps immediately based on the students’ responses.

Although discussed further in the next section, formative assessment is briefly defined in Figure 8.2.

Figure 8.2. What is Formative Assessment?

**What is formative assessment?** Formative assessment is a *process* teachers and students use *during* instruction that provides feedback to adjust ongoing teaching moves and learning tactics. It is *not* a tool or an event, nor a bank of test items or performance tasks. Well-supported by research evidence, it improves students’ learning in time to achieve intended instructional outcomes. Key features include:

1. **Clear lesson-learning goals and success criteria**, so students understand what they’re aiming for;
2. **Evidence of learning** gathered *during lessons* to determine where students are relative to goals;
3. **A pedagogical response to evidence, including descriptive feedback** that supports learning by helping students answer: *Where am I going? Where am I now? What are my next steps?*
4. **Peer- and self-assessment** to strengthen students’ learning, efficacy, confidence, and autonomy;
5. **A collaborative classroom culture** where students and teachers are partners in learning.

From Linn and Gage (2014, 2)

A second purpose of assessment is to provide information on students’ current levels of achievement after a period of learning has occurred. Such assessments—which may be classroom-based, districtwide, or statewide—serve a summative purpose and are sometimes referred to as assessments **of** learning. They help determine whether students have attained a certain level of competency after a more or less

extended period of instruction and learning, for example, at the end of a unit which may last several weeks, at the end of a quarter, or annually (National Research Council [NRC] 2001). Inferences made by teachers from the results of these assessments can be used to make decisions about student placement, instruction, curriculum, and interventions, and to assign grades. For example, the current state assessment of English language proficiency (ELP), the California English Language Development Text (CELDT), measures an EL's annual progress in attaining ELP. School districts use the results of the CELDT to make decisions about the ongoing instructional placement or possible reclassification of ELs. The CELDT will be replaced by the English Language Proficiency Assessments for California (ELPAC). (See page 60.)

As part of a balanced and comprehensive assessment system, assessment **for** learning and assessment **of** learning are both important. Another way to view this distinction is to note that assessment(s) of learning usually involve a tool or event *after* a period of learning, while assessment for learning is a process. Any evidence-gathering strategy used during formative assessment must yield information that is *timely* and *specific* enough to assist learning while it is occurring. Figure 8.3 highlights differences in key dimensions of these assessment purposes.

Figure 8.3. Key Dimensions of Assessment **for** Learning and Assessment **of** Learning

<b>Assessment:</b> <b>A Process of Reasoning from Evidence to Inform Teaching and Learning</b>			
<b>Dimension</b>	<b>Assessment <i>for</i> learning</b>	<b>Assessment <i>of</i> learning</b>	
<b>Method</b>	<b>Formative Assessment Process</b>	<b>Classroom Summative/ Interim/Benchmark Assessment*</b>	<b>Large-scale Summative Assessment</b>
<b>Main Purpose</b>	Assist immediate learning (in the moment)	Measure student achievement or progress (may also inform future teaching and learning)	Evaluate educational programs and measure multi-year progress
<b>Focus</b>	Teaching and learning	Measurement	Accountability
<b>Locus</b>	Individual student and classroom learning	Grade level/ department/school	School/district/state

<b>Priority for Instruction</b>	High	Medium	Low
<b>Proximity to learning</b>	In-the-midst	Middle-distance	Distant
<b>Timing</b>	<i>During</i> immediate instruction or sequence of lessons	<i>After</i> teaching-learning cycle → <i>between</i> units/periodic	<i>End</i> of year/course
<b>Participants</b>	Teacher and Student (T-S / S-S / Self)	Student (may later include T-S in conference)	Student

Adapted from Linquanti (2014)

\*Assessment of learning may also be used for formative purposes *if* assessment evidence is used to shape future instruction. Such assessments include weekly quizzes; curriculum embedded within-unit tasks (e.g., oral presentations, writing projects, portfolios) or end-of-unit/culminating tasks; monthly writing samples, reading assessments (e.g., oral reading observation, periodic foundational skills assessments); and student reflections/self-assessments (e.g., rubric self-rating).

As Figure 8.3 illustrates, assessment for learning—comprising key practices of the formative assessment process—occurs during instruction (or while learning is happening) and assists students' immediate learning needs. As it is intertwined and inseparable from teachers' pedagogical practice, formative assessment is of the highest priority. It is especially important in assessing and guiding students forward in developing the broad range of language and literacy skills and their application. Note also in Figure 8.3 that some assessments of learning can be used for formative purposes—that is, they can be used to inform future teaching and learning (and not simply to report on achievement or progress). This is only the case *if* the evidence gathering tool serves *both* the focus of instruction of the previous unit (or period of learning students just engaged in) *and* immediate future learning goals.

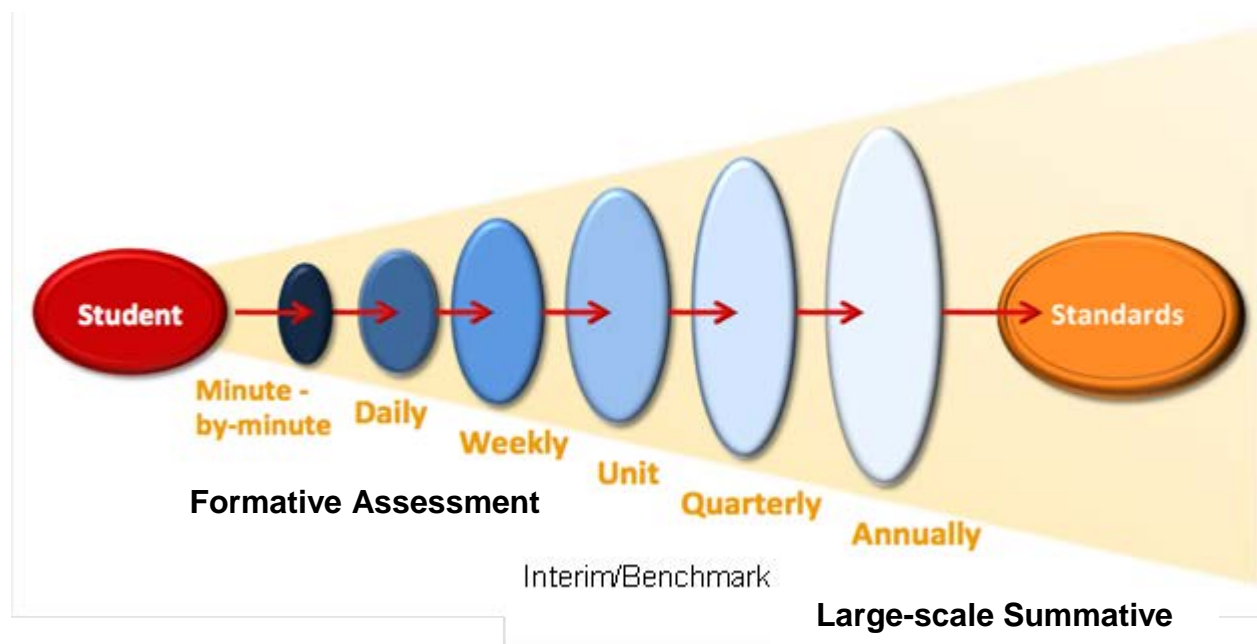
Classroom teachers, school leaders, and professional learning providers should consider the support that educators require in order to understand and implement the formative assessment process, as well as to use interim/benchmark and summative assessments effectively. Importantly, educators need to interpret assessment evidence in order to plan instruction and respond pedagogically to emerging student learning.

Collaborative professional environments, such as communities of practice, should be the nexus of learning and work that teachers do relative to assessment evidence as part of an ongoing cycle of inquiry (see Chapter 11). To maximize the use of assessment information for decisions related to student achievement of the CA CCSS for ELA/Literacy and the CA ELD Standards, teachers and leaders need to make full use of assessment for both formative and summative purposes.

### Assessment Cycles

One way to think about assessment for different purposes is to conceptualize assessment as operating in different cycles: short, medium, and long (William 2006). Figure 8.4 shows a range of assessments within a comprehensive assessment system. Those assessments that are more proximate to student learning (i.e., minute-by-minute, daily, weekly) operate in a short cycle because they address a short period of teaching and learning. Short-cycle assessment serves a formative purpose because its intended use is to inform immediate teaching and learning. Assessments administered at the end of the year are long-cycle because they cover a much longer period of learning. They are primarily used for summative purposes.

Figure 8.4. Assessment Cycles by Purpose



Adapted from Herman and Heritage (2007)

Occupying a middle position between short-cycle (formative) and annual (summative) assessment is interim/benchmark assessment: “assessments administered periodically throughout the school year, at specified times during a curriculum sequence to evaluate students’ knowledge and skills relative to an explicit set of longer-term learning goals” (Herman, Osmundson, and Dietel 2010, 1). In Figure 8.4, classroom summative assessments are referred to as unit assessments (although they could also occur in shorter time frames), and interim/benchmark assessments are referred to as quarterly assessments. Such periodic assessments operate in a medium cycle because they address longer-term goals than those assessments more proximate to student learning but not as long-term as annual assessments. Classroom summative or interim/benchmark assessments are generally used for summative purposes—evaluating what has been learned—although they may be used for formative purposes if they inform decisions that teachers and instructional leaders make within the school year regarding curriculum, instructional programs, professional learning, and so forth to improve future student learning. However, they are distinct from the formative assessment process because, by their design and intended use, they do not inform immediate teaching and learning. Unit assessments primarily serve a summative function but can serve a formative purpose if the teacher can act on the assessment information to support improved learning in a future unit. Progress-monitoring assessments can be short, medium, or long cycle, depending on whether they are administered after a longer or shorter period of instruction and they can serve both a formative and summative function. (For more information on screening, diagnostic assessment, and progress monitoring, see subsequent sections of this chapter).

Assessments within each cycle function best when they are part of a comprehensive, coherent and continuous system of assessment that provides ongoing information to teachers throughout the year (NRC 2001). Within such systems, minute-by-minute, daily, and weekly assessment feeds into unit assessment, which, in turn, feeds into periodic (e.g., end-of-unit, quarterly interim/benchmark) assessments, and multiple interim assessments feed into the annual assessment of the standards. A comprehensive, coherent and continuous system of assessment provides mutually complementary views of student learning, ensures that assessment within each cycle is

focused on the same ultimate goal—achievement of standards—and push instruction and learning in a common direction (Herman 2010).

Each assessment cycle provides information at varying levels of detail and inferences drawn from the assessment results are used to address specific questions about student learning and inform a range of decisions and actions. Figure 8.5 summarizes the types and purposes of the assessments within each assessment cycle.

Figure 8.5. Types and Uses of Assessments Within Assessment Cycles

Cycle	Methods	Information	Uses/Actions
<b>Short</b>			
<b>Minute-by-minute</b>	<ul style="list-style-type: none"> <li>-Observation</li> <li>-Questions (teachers and students)</li> <li>-Instructional tasks</li> <li>-Student discussions</li> <li>-Written work/representations</li> </ul>	<ul style="list-style-type: none"> <li>-Students' current learning status, relative difficulties and misunderstandings, emerging or partially formed ideas, full understanding</li> </ul>	<ul style="list-style-type: none"> <li>-Keep going, stop and find out more, provide oral feedback to individuals, adjust instructional moves in relation to student learning status (e.g., act on "teachable moments")</li> </ul>
<b>Daily Lesson</b>	Planned and placed strategically in the lesson: <ul style="list-style-type: none"> <li>-Observation</li> <li>-Questions (teachers and students)</li> <li>-Instructional tasks</li> <li>-Student discussions</li> <li>-Written work/representations</li> <li>-Student self-reflection (e.g., quick write)</li> </ul>	<ul style="list-style-type: none"> <li>-Students' current learning status, relative difficulties and misunderstandings, emerging or partially formed ideas, full understanding</li> </ul>	<ul style="list-style-type: none"> <li>-Continue with planned instruction</li> <li>-Instructional adjustments in this or the next lesson</li> <li>-Find out more</li> <li>-Feedback to class or individual students (oral or written)</li> </ul>
<b>Week</b>	<ul style="list-style-type: none"> <li>-Student discussions and work products</li> <li>-Student self-reflection (e.g., journaling)</li> </ul>	<ul style="list-style-type: none"> <li>-Students' current learning status relative to lesson learning goals (e.g., have students met the goal(s), are they nearly there?)</li> </ul>	<ul style="list-style-type: none"> <li>-Instructional planning for start of new week</li> <li>-Feedback to students (oral or written)</li> </ul>

Cycle	Methods	Information	Uses/Actions
<b>Medium</b>			
<b>End-of-Unit/ Project</b>	<ul style="list-style-type: none"> <li>-Student work artifacts (e.g., portfolio, writing project, oral presentation)</li> <li>-Use of rubrics</li> <li>-Student self-reflection (e.g., short survey)</li> <li>-Other classroom summative assessments designed by teacher(s)</li> </ul>	<ul style="list-style-type: none"> <li>-Status of student learning relative to unit learning goals-</li> </ul>	<ul style="list-style-type: none"> <li>-Grading</li> <li>-Reporting</li> <li>-Teacher reflection on effectiveness of planning and instruction</li> <li>-Teacher grade level/departmental discussions of student work</li> </ul>
<b>Quarterly/ Interim/ Benchmark</b>	<ul style="list-style-type: none"> <li>-Portfolio</li> <li>-Oral reading observation</li> <li>-Test</li> </ul>	<ul style="list-style-type: none"> <li>-Status of achievement of intermediate goals toward meeting standards (results aggregated and disaggregated)</li> </ul>	<ul style="list-style-type: none"> <li>-Making within-year instructional decisions.</li> <li>-Monitoring, reporting; grading; same-year adjustments to curriculum programs</li> <li>-Teacher reflection on effectiveness of planning and instruction</li> <li>-Readjusting professional learning priorities and resource decisions</li> </ul>
<b>Long</b>			
<b>Annual</b>	<ul style="list-style-type: none"> <li>-Smarter Balanced Summative Assessment</li> <li>-CELDT</li> <li>-Portfolio</li> <li>-District/school created test</li> </ul>	<ul style="list-style-type: none"> <li>Status of student achievement with respect to standards (results aggregated and disaggregated)</li> </ul>	<ul style="list-style-type: none"> <li>-Judging students' overall learning</li> <li>-Gauging student, school, district, and state year-to-year progress</li> <li>-Monitoring, reporting and accountability</li> <li>- Classification and placement (e.g., ELs)</li> <li>-Certification</li> <li>-Adjustments to following year's instruction, curriculum, programs;</li> <li>-Final grades</li> <li>-Professional learning prioritization and resource</li> </ul>

Cycle	Methods	Information	Uses/Actions
			decisions -Teacher reflection (individual/grade level/department) on overall effectiveness of planning and instruction

### Short-Cycle Formative Assessment

Short-cycle formative assessment is a process used by teachers and students *during instruction* that provides feedback to adjust ongoing teaching and learning to improve student achievement of intended instructional outcomes (McManus 2008). Short-cycle formative assessment occurs when evidence of learning is gathered minute-by-minute, daily, and weekly from a variety of sources during ongoing instruction for the purpose of moving learning forward to meet short-term goals (i.e., lesson goals) (Black and Wiliam 1998; Council of Chief State School Officers Formative Assessment State Collaborative 2006; Heritage 2010; Popham 2010). In the remainder of this chapter, this short-cycle formative assessment process is referred to as formative assessment.

This type of assessment provides the most detailed information for teachers and their students. The idea of formative assessment, or assessment *for* learning, does not apply to a specific tool or assessment. This is not to say that a tool or assessment cannot be used for formative assessment purposes—it can, but only if it provides actionable information about students’ learning status relative to the desired lesson goal and teachers can use it immediately to adjust their instruction. Many assessments marketed under the formative assessment label do not provide information needed about students’ learning in order to adjust instruction and guide students’ learning while it is occurring (Perie, Marion, and Gong 2009; Shepard 2005).

The sources of evidence available to teachers in short-cycle formative assessment are what students do, say, make, or write (Griffin 2007). For example, sources of evidence can be teacher-student interactions fuelled by well-designed questions (Bailey and Heritage 2008; Black, and others 2003), structured peer-to-peer discussions that the teacher observes (Harlen 2007), dialogues that embed assessment into an activity already occurring in the classroom (Ruiz-Primo and Furtak 2004, 2006

2007), student work from well-designed tasks (Poppers 2011), and web-based reading assessments that provide immediate feedback (Cohen, and others 2011).

The report of the Formative Assessment for Students and Teachers/State Collaborative on Assessment and Student Standards (FAST/SCASS) Project of the Council of Chief State School Officers (CCSSO) emphasizes several features of formative assessment. First, “formative assessment is a *process* rather than a particular kind of assessment.... There is no such thing as a ‘formative test’” (McManus 2008, 3). Second, “the formative assessment process involves both teachers *and* students..., both of whom must be actively involved in the process of improving learning” (3). Third, teachers must be clear about the ultimate goal of a unit and the sub-goals or stepping stones that are important along the way: “...from a learning progression teachers have the big picture of what students need to learn, as well as sufficient detail for planning instruction to meet short-term goals” (4). Fourth, success criteria and evidence of learning need to be laid out at the beginning of the project and reviewed along the way: “...teachers must provide the criteria by which learning will be assessed... using language readily understood by students, with realistic examples of what meets and does not meet the criteria.”

Whatever the source of the evidence, the teachers’ role is to construct or devise ways to elicit responses from students that reveal where they are in their learning and to use the evidence to move learning forward (Sadler 1989). For effective formative assessment, teachers need to be clear about the short-term learning goals (for example, for a lesson) that cumulatively lead to students’ attainment of one or more standards. They will also need to be clear about the success criteria for the lesson goal—how will the students show if they have met, or are on the way to meeting the lesson goal. The evidence-gathering strategy can then be aligned to the success criteria.

Questions that formative assessment can answer include the following:

- Where are my students in relation to learning goals for this lesson?
- What is the gap<sup>2</sup> between students' current learning and the goal?
- What individual difficulties are my students having?
- Are there any missing building blocks in their learning?
- What do I need to adjust in my teaching to ensure students learn?

Information from formative assessment is used to make instructional adjustments in real time, to continue with the planned lesson, or to provide feedback to students that will help them take steps to advance their learning. (Feedback to students is discussed in the Student Involvement section of this chapter.)

An important point about teachers' use of evidence in formative assessment is that their inferences from the evidence and their actions in response to that evidence focus on individual students. This does not mean that instruction for students is necessarily on a one-to-one basis, but rather that individual needs are addressed in the context of a class of students. This orientation to individuals is necessary if students are going to have the opportunity to learn and progress equally (Heritage 2013). To do so, instruction needs to be contingent upon each student's current learning status. In other words, instruction has to be matched to where the students are so that they can be assisted to progress and meet desired goals.

While formative assessment evidence is not aggregated in the same way as medium- and long-cycle assessment information, teachers can categorize individual student responses to look for patterns across the class or for particular students who are outliers. For example, at the end of a lesson after students have completed a response to a question about a text, a teacher can quickly categorize them into students who are showing they understand, students who are nearly there, and students who need more work. The next day's instruction can be planned accordingly.

---

<sup>2</sup> The gap refers to the distance between where the students' learning currently stands at particular points in the lesson (a lesson can be several periods or days long) and the intended learning goal for the lesson. The purpose of short-cycle formative assessment is to close this gap so that all students meet the goal (Sadler 1989). This should not be confused with the term *achievement gap*, which refers to differences in summative educational outcomes among different subgroups of students.

Teachers of ELs should take great care in making these formative assessment decisions. Depending on their level of English language proficiency, some ELs may not be able to fully express their ideas orally about a topic during a class discussion; however, this does not necessarily mean that they do not understand the topic. In addition, an informal observation that indicates that ELs are not orally proficient in English should not determine how the students should be taught reading in English. English learners do not have to be proficient in oral English before they can learn to read in English (Bunch, Kibler, and Pimental 2012). Teachers should use a combination of observations (e.g., during collaborative conversations between students about texts read) and informal inventories of reading (e.g., listening to students read aloud during small reading group time, asking specific comprehension questions to elicit student understandings) in order to determine how best to support their ELs and provide *just-in-time* scaffolding as they progress in their reading development. In addition, the CA ELD Standards clearly specify that all ELs, regardless of their level of English language proficiency, are capable of engaging in intellectually-rich tasks at the same cognitive level as their English-proficient peers. To help ensure this happens, teachers can use in-the-moment formative assessment practices to provide the appropriate level of scaffolding for ELs. (For more information on scaffolding, see Chapter 2.)

Using the formative assessment process in an EL student's primary language, in contexts where teaching and learning utilize this resource, such as in an alternative program, may also offer instructionally actionable information. For example, a newcomer EL at the Emerging level of English language proficiency (e.g., a student who has been in the U.S. for less than a year) may find it difficult to respond (in writing or orally) to a question about their knowledge of a science or history topic in English with the same level of detail as they might be able to do in their primary language. Teachers could ask their newcomer EL students to quickly write responses to text-based questions first in their primary language (if they are literate in it) before they respond in written English. Then, the two pieces of writing could be compared in order to identify similarities and differences between both content knowledge and literacy in the primary language and English. This technique could be applied strategically so that teachers ensure they have a clear understanding of what students know about particular topics

and how well they are able to express their understandings in English. Teachers can also use this type of evidence to explicitly draw their EL students' attention to ways they could express through English writing or speaking what they already know and are able to convey in their primary language. While all teachers may not be able to provide this type of support themselves (e.g., when they are not proficient in their EL students' primary languages), they may be able to collaborate with other teachers, EL specialists, or community members to do so.

The use of technology that enables students to give immediate responses to teachers (e.g., clickers, mobile devices) can also help teachers with large numbers of students to get an ongoing sense of where students are during the lesson. For example, halfway through a lesson, a tenth grade teacher asks three or four questions related to multiple meanings and word phrases in a literary text the class has been analyzing. The results immediately appear as a pie chart on the smart board. Both teachers and students can quickly see how the class responded and can decide together if more work needs to be done in this area before the lesson progresses.

The following snapshots provide additional concrete examples of formative assessment in action.

#### **Snapshot 8.1 Formative Assessment in Grade Five**

Fifth graders are working on the following CA CCSS for ELA/Literacy standards: a) applying the reading standard for informational text: *explaining how an author uses reasons and evidence to support particular points in a text, identifying which reasons and evidence support which particular points* (RI.5.8); b) the writing standard: *produce clear and coherent writing in which the development and organization are appropriate to task, purpose, and audience* (W.5.4); and the language standard *vocabulary use* (L.5.4-6), particularly transition words to help their writing flow logically. They are writing an argument to encourage their readers to take more care of the natural environment. In their reading instruction, they have analyzed a text to identify where the sections with the *arguments*, *counterarguments*, and *evidence* to support the arguments are located. In their writing, they are learning to apply these ideas, as well as how to organize their arguments effectively.

While the students are involved in the independent writing part of the lesson, the teacher sits with a student to discuss his writing progress. She has a ring binder open to a page with these headings at the top: *Child's Name/Date*, *Research Compliment*, *Teaching Point*, and *What's Next for this Child?* Further down the page is a self-adhesive note that lists five students' names. These are the other children she

plans to meet with during the session.

The teacher's initial purpose with the student is to follow up from two days ago when she provided him with feedback based on the evidence she had elicited from her interaction with him; in that interaction she determined that the student needed to provide stronger sources of evidence to support his argument. On this occasion, she wants to see how he has used her prior feedback:

T: *You're working on evidence? Tell me about it.*

S: *I found good information in the book of the Environmental Protection Agency and on the Internet.*

T: *And what do you think about what you found so far? Do you think that it supports your argument?*

S: *I guess....*

At this stage, the teacher reminds the student that the purpose of the evidence is to support his argument. She explains what "supporting an argument" is, in a way that is meaningful to a fifth grader, by telling him that you have to prove it with what is in the text or the readers may not believe you. She asks him to read his argument aloud. Having established that the focus of his argument is to "stop dumping in the ocean because all the beautiful animals we see are going to start vanishing," the teacher asks: *So, what evidence did you find to support that claim—that all the animals will die if we don't stop dumping? What evidence did you find that will help you to strengthen that argument, or prove it to your readers?* The teacher then helps her student to recognize which of the information he has located is from a reliable source and will be effective in supporting his argument. Satisfied that the student can move forward on his own to incorporate his evidence, she then asks him to go over the organization of his argument and to let her know where he will place the evidence. When the student does this, it is evident to the teacher that he has some confusion about the overall structure and that his writing needs to be reorganized. This is a moment in the interaction when she targets a teaching point for him. She goes over the organization with him and writes the organizational elements on a self-adhesive note and includes specific support, such as putting the evidence in order to help the flow or adding transitional sentences.

Throughout this interaction, the teacher has made notes in her ring-binder file. Under *Research Compliment* she has written that the student recognized the reliability of his source, in the section labeled *Teaching Point* she wrote that she had discussed how evidence supported his argument, and under the heading *What's Next for this Child?* she wrote "organization and transitional sentences," noting that the student had problems organizing his writing to effectively convey his argument to the reader. By gathering evidence in the course of this interaction, the teacher was able to match her teaching points to the specific student's needs. Additionally, after several interactions of this kind, she may find that there are common needs among several students and might pull them together for a mini-lesson.

**Snapshot 8.2 Formative Assessment in Grade Two**

In a second-grade classroom comprised of native English speaking children and children who are ELs, the children have been working on retelling folktales they have read together in class and conveying the central message of the tale (RL.2.2). The EL children, in particular, have been working on using the past tense to indicate the tales happened in the past (ELD.PII.2.3). In this lesson, students are engaged in small group work, and during this time, the teacher selects groups of three students to recount one of the folktales the class has read that week. In this situation, she wants to give each student sustained opportunities to use language while she and the others in the group listen. She asks the first student to begin, then after a while asks the second child to carry on and so on. When the students have finished, the teacher asks them to say what they think the main message of the story is. Each child offers an opinion and there follows a discussion about whether there is agreement about the main message or not. From the recounting activity, the teacher has evidence that one student uses the past tense consistently and mostly with accuracy, while the other two do not. Two of the children are able to convey the message of the text, but another has not really grasped it. After her discussion with the group, she makes quick notes about each student and what is next for them instructionally. She continues this process with one more group before independent reading time is over, and she will find other opportunities during the week to assess other small groups in the same way.

**Snapshot 8.3 Formative Assessment with Secondary EL Newcomers**

In a secondary designated English Language Development (ELD) class, with newcomers whose experience in the U.S. ranges from three months to one year, the ELD teacher has worked collaboratively with the science teacher to create a five-week unit on animal behavior with the purpose of guiding his students through a deep exploration of the content through the language resources used to convey meaning. The two teachers have agreed that during science instruction, the science teacher will provide appropriate and strategic support so to his EL students so that they can fully participate in the science activities he has designed, gain understanding from the science textbook, and engage in collaborative discussions about the text and content. This strategic support includes using graphic organizers, providing increased opportunities for the students to discuss their ideas in small groups or pairs, and primary language support, including drawing attention to cognates and using texts in students' primary languages.

During designated ELD instruction, the ELD teacher has agreed to analyze the science textbook and the activities the science teacher has designed in order to identify the language demands they present and then to address the language demands more intensively in her class. This is the third class of the first week on the unit. Having formulated questions they would like to explore around the science topic, students have then perused a variety of texts on the topic to identify meanings and have charted language (including phrasing and general academic and domain-specific vocabulary) they think is critical for conveying their understanding of the topic. They now work in pairs to collaboratively write a

description about what they have learned so far about one aspect of animal behavior, using as much of the language they have charted as they can. Before the lesson is over, the pairs write their description drafts on large sheets of paper to enable a discussion on what they have done and where they may go next to refine or add to their descriptions. The pairs read their descriptions to the class, and time is provided for the other students to ask questions and make comments. When one pair shares their description about animals and language, an animated conversation develops on whether animals have language or not. Julio, explains the thinking that went into the description that caused the lively discussion.

*Julio: ...First of all, I think that language is a way to **inform** others around you, your feelings or just a simple thing that you want to let know people what is the deal. And it can be **expressed** by saying it, watching a picture, or hearing it, you know what I'm saying? I don't know if you have heard about the kangaroo rat that stamps its feet to **communicate** with other rats. It's really funny cause we humans have more **characteristics** to **communicate** to each other, but we still have problems to understand other people. Characteristics like sound, grammar, pitch, and body language are some of them, while the rat only uses the foot (he stamps the ground).*

The teacher, who has been taking notes on the language students are using in the conversation, also notes that Julio is using some of the academic language the class has charted in both his writing and speaking and has, more importantly, done an effective job of conveying his understanding of the information from his research and persuading his peers using evidence. The ELD teacher decides to examine more closely the students' written descriptions, as well as the language they have used in their conversations, in order to make decisions about what language features of the science texts to focus on as she progresses in the unit. She also plans to make a copy of her notes to share with the science teacher when they meet later that week during collaboration time.

### **Medium-Cycle Assessment**

Assessments that teachers develop, or that are included in the curricular materials and are administered at the end of a unit, are medium cycle. As noted previously, medium-cycle assessments occupy a middle ground between short-cycle formative assessment and long-cycle summative assessments. Some are used to inform instruction during the school year; others serve evaluative purposes.

### **End-of-Unit Assessments**

End-of-unit assessments can serve a summative purpose to evaluate student achievement with respect to the goals of the unit. If such assessments are given to students before the end of the unit when there is still time to take some instructional action before moving on to the next unit, then they can also serve a formative purpose.

In developing unit assessments, teachers will need to ensure that the goals of the unit are clear and aligned to standards. In other words, what is to be assessed must be well articulated and derived specifically from the standards and lesson planning. When teachers know *what* to assess, they can best determine *how* to assess. In other words, they can decide on the most effective way that students can demonstrate the achievement of the goals.

End-of-unit assessments can help teachers answer questions such as the following:

- Have my students met the goals of the unit?
- Are there some students who need additional help to meet the goals of the unit?
- What help do they need?
- What improvements do I need to make in my teaching next time I teach this unit?

The following snapshot provides a concrete example of the use of end-of-unit (medium cycle) assessment.

#### **Snapshot 8.4 End-of-Unit (Medium-Cycle) Assessment in Grade Seven**

In a seventh grade classroom with native English speakers, recently reclassified ELs, and a group of ELs who are at the Expanding and Bridging levels of English language proficiency, the teacher has been taking the students through a five week unit: Persuasion Across Time and Space: Analyzing and Producing Complex Texts (*Understanding Language* 2013). This unit addresses multiple CA CCSS for ELA/Literacy and CA ELD standards simultaneously. The unit has four primary goals: 1) to read and analyze complex texts; 2) to involve students in reading, writing, listening, and speaking activities that are grounded in evidence from informational texts; 3) to engage students in disciplinary practices highlighting language and purpose that are responsive to audience; and 4) to build history/social studies knowledge through content rich non-fiction.

During the course of the unit, with intentional and strategic scaffolding by the teacher and considerable involvement in collaborative groups, the students engaged in close reading, collaborative discussions, and analysis of the text organization, grammatical structures, and vocabulary of persuasive texts on relevant topics. In the final part of the unit, the students analyzed the video, “The Girl Who Silenced the World for Five Minutes,” compared and contrasted persuasive techniques in the video to one of the texts they had read, and produced a persuasive text of their own. The students’ analysis of the video and written work served as the summative assessment for the unit. Using the students’ work, the teacher was able to make a determination about the students’ understanding of the purpose, organization, and structure of persuasive texts and their ability to use various language resources

(including vocabulary, complex grammatical structures, connecting words and phrases) to write a coherent and cohesive persuasive piece for a public audience.

After reviewing the students' responses, the teacher concluded that the students had made good progress toward meeting the goals of the unit, especially in regard to their understanding of persuasive techniques in different contexts (i.e., video and text). Examining her EL students' writing more closely, however, the teacher noticed that most of her students' writing was characterized by text that appeared more like spoken, everyday language. In other words, their written arguments were not making use of connecting words and phrases (e.g., for example, therefore, consequently) to create cohesion, nor were they using many complex sentences to connect ideas and create relationships between them (e.g., *Even though governments are taking action*, it is not happening fast enough). This analysis of her students' writing helped the teacher to design lessons in the very next unit where she could show them examples of cohesion and complex sentences that connect ideas, model how to *unpack* the meaning in the texts, collaboratively construct similar writing with the students, and provide them with guided practice in writing related to the unit topic. She also planned to draw her students' attention to various examples of persuasive language used in arguments and to observe how her students incorporated them into their own writing in the next units she had planned. In addition, she made a note to address these linguistic features directly when she teaches the unit the following year. (Snapshot adapted from *Understanding Language* 2013)

### ***Interim or Benchmark Assessments***

Interim or benchmark assessments, such as the Smarter Balanced interim assessments, are medium-cycle and address intermediate goals on the way to meeting standards. The Smarter Balanced assessments are aligned to the standards, and any other interim or benchmark assessment used by districts or schools will also need to be aligned to the standards. Typically administered quarterly or every six weeks, interim assessments cover a shorter period of instruction than long-cycle assessments and consequently give more detail about student learning. Results from interim assessments provide periodic snapshots of student learning throughout the year. These snapshots assist teachers to monitor how student learning is progressing and to determine who is on track to meet the standards by the end of the year and who is not, which may mean that a student is in need of additional support. When using or designing interim or benchmark assessments, teachers and school and district leaders should consider what is reasonable to expect students to be able to do with regard to *meeting* a CA CCSS for ELA/Literacy or CA ELD Standard at various points along the year (as opposed to at the

end of the year). While there should be goals for meeting the end-of-year standards at points along the school year, these goals may look different than the end of year standard. Results from these assessments can help teachers answer the following questions:

- What have my students learned so far?
- Who has and who has not met intermediate goals?
- Who is and who is not on track to meet the standards by the end of the year?
- How are students performing on this test or assignment on those areas identified as weak on last year's California state long-cycle assessments?
- What are the strengths and areas of need in an individual's or groups' learning?
- Who are the students most in need of additional support? What do they need?
- What are the strengths and areas of need in my curriculum?
- What are the strengths and areas of need in my instruction?
- What improvements do I need to make in my teaching?

Administrators can also use interim assessments to address many of these questions that are relevant to their decision-making needs, for example, programmatic, professional learning, and resource decisions.

If students are not making desired progress, then teachers and administrators should consider if changes are needed in curriculum and instruction while adjustments can still be made before the end of the year. In this sense, even though they sum up a period of learning (over a few weeks or months) their use is also *formative* if adjustments to curriculum and instruction are made during the school year. Interim assessments also supply individual performance data. These data are useful to identify individual student's strengths and learning needs. In addition, while these results sum up a period of learning, they can also be used formatively if steps are taken to respond to individual student's needs while there is still time within the year. In instances where no action is taken to support student learning, the results from these assessments remain summative only.

Using data systems, including spreadsheets, interim assessment results can be aggregated and displayed in graphs and charts, so teachers can identify patterns in their students' performance, and disaggregated to provide information on the relative

performance of individuals and subgroups. It is important to ensure that teachers and administrators have adequate training and support to properly interpret the results of interim assessments so their conclusions and responses will be appropriate and effective.

If districts, schools, or individual teachers use commercially-produced interim assessments, they must consider technical quality to ensure that the assessments are appropriate for the intended purpose and that they are fully aligned with the CA CCSS for ELA/Literacy and CA ELD Standards that comprise the learning goals of the period of learning being assessed. (See section on Technical Quality in this chapter.)

#### **Snapshot 8.5 Interim ( Medium-Cycle) Assessment in Grade One**

All incoming first graders in a school are assessed at the beginning of the school year on the foundational skills of the CA CCSS for ELA Standards, specifically, print concepts, phonological awareness, phonics and word recognition, and fluency. Results from their end-of-year kindergarten assessment are used to determine which sections of the assessment they receive. For example, if a student's results indicated a complete understanding of print concepts, that part of the assessment would be skipped, although close observations would be made during class to confirm last year's assessments. The teachers find the results from the beginning of the year assessment to be a useful starting point for their instructional planning, particularly as students may have either lost or made up ground during the summer. In addition, the teachers assess, or obtain help to assess, the primary language foundational literacy skills of their ELs who are new to the school and use this information for instructional decision-making.

After these initial assessments and appropriately designed instruction, students are administered interim foundational skills assessments every six weeks to determine progress. While the teachers are regularly using formative assessment practices during their instruction to gather evidence of students' skill development and to adjust instruction accordingly, they use the results of the interim assessments to gauge the overall progress of individuals and the class as a whole, and to indicate to them where they need to make improvements in their teaching to ensure greater progress. The teachers also use the results as a means to evaluate and corroborate their own judgments about students' skill development in the period between the interim assessments' administration.

#### ***Assessing ELD Using Medium-Cycle Evidence***

As with all medium-cycle assessment, assessing progress in English language development using interim/benchmark assessments or classroom summative (such as

end-of-unit/quarter) assessments should not take priority over short-cycle formative assessment. Short-cycle ELD formative assessment is assessment *for* learning and allows immediate teaching moves that support student learning *as it is occurring*. Medium-cycle ELD assessment is assessment *of* learning that has already occurred and is typically not useful for providing immediate instructional support to students because the assessment evidence is too removed in time from the learning and often too general to be useful *for* learning.

This is not to say that medium-cycle ELD assessment is not important. On the contrary, classroom summative and interim/benchmark assessment are useful for evaluating a student's progress. They can help teachers reflect on their instructional planning and implementation and make within-year program design and instructional adjustments. They can also help school and district leadership identify or adjust professional learning and instructional materials decisions. This periodic progress monitoring helps teachers determine the status of EL students' achievement of unit and intermediate (e.g., within-unit, quarterly) goals toward meeting particular CA ELD Standards as they progress through English language proficiency (ELP) levels. In addition, they help teachers determine if students are in fact advancing in English language proficiency (ELP) or if they seem to be *stalled* in particular areas. In other words, they help teachers know if their students are *on track* for achieving end-of-year learning goals, which are differentiated for ELs using the CA ELD Standards as a guide, so that within-year adjustments and refinements can be made to instruction. For example, a fifth grade teacher could examine a quarterly narrative writing task (e.g., a story) the whole class completes and use the CA ELD Standards to analyze how her EL students expand and enrich their ideas in noun phrases (ELD.PII.5.4). Using this approach, she is able to monitor how an individual EL student progresses throughout the year. A student who began the school year at an early Expanding level of ELP, for example, might progress through the Expanding level (across the story writing samples) and, potentially, into the early Bridging level by the end of the year.

Similarly, a high school teacher may design a two-month unit of study with a culminating, curriculum-embedded argument writing task. This writing task can be a very useful source of medium-cycle ELD assessment evidence if the writing is analyzed

for degree of attainment in the learning goals tied to particular CA ELD Standards (how students are using verb types and tenses, organizing their writing, expanding noun phrases, etc.). This analysis helps teachers identify patterns in student learning outcomes (e.g., many students may need support in linking ideas throughout a text to create cohesion) and individual student areas for improvement. If teachers use the results of this analysis in planning further instruction (which they should), this within-year adjustment supports students to progress in ELD toward end-of-year goals.

Interim/benchmark assessments should be used judiciously and intentionally. The best way to inform ongoing teaching and learning regarding how EL students interact meaningfully in English and learn about how English works is through authentic classroom learning tasks, and not via multiple-choice tests or decontextualized performance tasks that focus on discrete grammatical skills and vocabulary knowledge. Teachers should approach these types of assessments with a critical eye, ensuring that the tests match teaching and learning goals and that valuable instructional time is not sacrificed for administering and analyzing tests—or any other type of medium-cycle assessment—that are not clearly critical for monitoring ELD progress.

### **Long-Cycle Assessment**

Yearly assessments, such as the Smarter Balanced annual assessments, are long-cycle assessments. They cover a year's worth of learning and, by their nature, provide very large grain-sized information about student achievement relative to the standards. They sum up achievement after a year of learning and are therefore most appropriately used by schools and districts to monitor their own annual and longitudinal progress and to ensure individual students are on track academically. The CELDT serves similar purposes with respect to measuring ELs' progress toward learning English. Schools and districts can ensure that students in dual language programs are making steady progress toward biliteracy by including assessments in the relevant non-English language.

Long-cycle assessments are also useful to teachers and can help them answer such questions as:

- What did my outgoing class of students learn? Did they meet the standards I was teaching them?

- What did my incoming class of students learn from last year to this year? Which standards did they achieve, and which did they not achieve?
- What are the overall strengths and areas of need in my class's learning?
- What are the strengths and areas of need in individual's and groups' learning?
- What are the strengths and areas of need in my curriculum?
- What are the strengths and areas of need in my instruction?
- Have the improvement strategies I/we put in place worked?

With data systems, the assessment results can be aggregated so that individual teachers and schools can look for patterns in their students' performance. They can also be disaggregated to provide information on the relative performance of subgroups and the performance of individual students. School and district administrators can also use these assessment results to address questions relating to which students have and have not met the standards, and the relative strengths and areas of need in curricula and programs. Successful schools discuss long-cycle assessment with an eye to responsively adjusting the ways they analyze outcomes, plan professional learning, collaborate, and teach.

Long-cycle assessment results are appropriately used for system monitoring and accountability, reporting to parents on their individual child's achievement, adjustments to programs, curriculum and instruction for the following school year, teachers' reflection on their instructional practices, and identifying teachers' professional learning needs. The results also provide a starting point for the students' teachers the following school year, in terms of a picture of a class's, a subgroup's and an individual's strengths and weaknesses. Snapshot 8.6 provides a glimpse of these uses of long-cycle assessment.

#### **Snapshot 8.6 Long-Cycle Assessment in Grade Eight**

During the summer of the new school year, just before school starts, eighth grade English teacher Ms. Flora and her eighth-grade colleagues examine their incoming eighth graders' seventh grade summative ELA assessment results in order to anticipate their students' learning needs. At the same time, they examine the prior year's CELDT results for their EL students, some of whom have been in U.S. schools for only a couple of years and others for many years, as well as available data about their literacy proficiency in their primary language. The teachers want to make sure that they use all available information to design appropriately differentiated instruction for each of their students.

As a result of analyzing last year's results, which suggested students may need much support in particular areas, the teachers work diligently to improve the students' close and analytic reading skills with respect to literature and informational text, and their ability to write arguments effectively. To address weaknesses evident in the seventh grade summative assessment results, Ms. Flora pays particular attention to the eighth grade literature standards: 1) Cite textual evidence that most strongly supports an analysis of what the text says explicitly as well as inferences drawn for the text (RI.7.1), and 2) Compare and contrast the structure of two or more texts and analyze how the differing structures of each text contribute to its meaning and style (RL.8.5). She parallels the focus of the first literature standards in informational text as well. In addition, to address the weaknesses evident in the seventh grade writing results, she works with her students extensively on the standard: Write arguments to support claims with clear reasons and relevant evidence (W.6-8.1).

The following year, when she examines her students' eighth grade ELA summative assessment results, the first question she wants to answer is: Have my students met the standards I was focusing on with particular intensiveness? She notes that most students have achieved proficiency with respect to the reading and writing standards. She is satisfied with the overall result and feels that the instructional focus that she and her colleagues identified for the year has yielded positive results. However, there are more students who do not meet the proficient levels on the state achievement assessment than she would like, so she plans to follow up with her colleagues to look at the overall grade level performance to identify if there are students in other classes that are in these categories. She also plans to investigate the scores of individual students who have not met the standard to see where specific areas of need lie and if the results of summative assessment are consistent with what she has observed through formative assessment and interim assessments. For her EL students, she will compare all of this information to their eighth grade summative CELDT assessment results and note any relevant findings.. This information provides evidence to help guide any changes in her instruction for next year's eighth graders. She also knows that her careful analyses will be valuable information to pass on to the ninth grade teachers.

### ***Ensuring Accessibility for ELs on Long-Cycle Assessments***

To ensure as accurate a picture as possible of ELs' learning status, designated assessment supports may be needed. The intent is not to give EL students an unfair advantage over those who are not receiving that support (Abedi and Ewers 2013). The goal of a support is to make an assessment more accessible for ELs, and to allow them to demonstrate what they know and can do, thereby leveling the playing field and strengthening the validity of inferences from assessment results for these students. There are five major considerations when selecting assessment supports for ELs:

- Effectiveness: a support must be effective in making an assessment more

accessible to the recipients;

- Validity: a support should not alter the focal construct being assessed, that is, the outcomes of supported and unsupported assessments should be comparable;
- Differential Impact: a support should be sensitive to students' background characteristics and their academic standing; one size may not fit all;
- Relevance: a support should be appropriate for the recipients;
- Feasibility: a support must be logistically feasible to implement in the assessment setting (Abedi and Ewers 2013, 4).

The Smarter Balanced assessment consortium offers universal embedded online tools that improve the accessibility for all students, several embedded designated supports that improve accessibility for ELs, and accommodations for those students with disabilities that require them under their individualized education program (IEP) or 504 plan (Smarter Balanced 2013b). Examples of designated supports, depending on the type of assessment, include bilingual glossaries, translated test directions, and text-to-speech features. Because the type of support useful to ELs varies depending on the student's age, the student's level of English language proficiency, the subject area assessed, type of assessment task, and other factors, supports should be used strategically and intentionally. (See Usability, Accessibility, and Accommodations Guidelines from Smarter Balanced 2014 [http://www.smarterbalanced.org/wordpress/wp-content/uploads/2013/09/SmarterBalanced\\_Guidelines\\_091113.pdf](http://www.smarterbalanced.org/wordpress/wp-content/uploads/2013/09/SmarterBalanced_Guidelines_091113.pdf) ).

### **Additional Methods of Medium- and Long-Cycle Assessment**

Additional methods for evaluating student achievement in medium or long cycles include rubrics and student portfolios.

#### ***Rubrics***

Performance assessments that require students to demonstrate learning through an oral, written, or multimodal performance task (e.g., a presentation, a report) are can be evaluated according to a rubric. A commonly accepted definition of a rubric is that of a document that articulates the expectations for an assignment by listing the criteria, or what counts, and describing levels of quality (Andrade, and others 2009). Criteria should relate to the learning that students are being asked to demonstrate rather than

the tasks themselves, and they should provide clear descriptions of performance across a continuum of quality (Brookhart 2013). The criteria should be linked to standards and reflect what is required to meet a specific standard or cluster of standards.

Descriptions of performance are usually presented within score levels, and the number of score levels depends on the extent to which criteria across the levels can distinguish among varying degrees of understanding and skills. The knowledge and skills at one level should differ distinctively from those at other levels (Lane 2013). If schools are using commercially produced performance assessments for high stakes assessment purposes, for example placement or end-of-year grades, they will need to ensure that the rubrics have undergone a series of studies that provide evidence of their technical quality. (See the section on Technical Quality in this chapter.) Examples of such studies include review by language and literacy experts, review to ensure cultural and language sensitivity, and field tests to provide evidence that the rubric can differentiate performance across levels of the rubric and across grades.

For classroom assessment, in those situations where stakes are not so high, teachers—sometimes in collaboration with students—can also develop rubrics for their own classroom performance assessments. Co-construction of rubrics with students is a powerful way to build student understanding and acceptance of what is expected. When creating rubrics, there are a few points to bear in mind. First, rubrics should express as clearly and concisely as possible the expected performance at each level. Therefore, it is important to avoid ambiguous language. Before using the rubric, the language of the rubric will need to be explained to students. Second, in communicating expectations non-pejorative descriptions of what performance looks like at each level, reflecting a growth mindset, should be used. Third, the gradations of quality need to be specifically articulated across levels. Figure 8.6 shows an example of a rubric for scoring an essay. The dimensions of the rubric are listed on the left-hand side and the criteria are clearly described across four levels of performance.

Figure 8.6. Essay Scoring Rubric

Dimensions	4	3	2	1
<b>Ideas and Content</b>	The essay has a clear thesis and supports it with evidence. Relevant comparisons b/w the paintings are made. Reasons for the similarities and differences are discussed in terms of the influence of one art movement on another.	The essay has a clear thesis. Comparisons b/w the art works are made. The discussion of influences might be thin.	An opinion is given. The support for it tends to be weak or inaccurate. May get off topic.	The thesis and support for it is buried, confused and/or unclear.
<b>Organization</b>	The paper has an interesting beginning, developed middle, and satisfying conclusion in an order that makes sense. Paragraphs are indented, have topic and closing sentences, and main ideas.	The paper has a beginning, middle and end in an order that makes sense. Paragraphs are indented; some have topic and closing sentences.	The paper has an attempt at a beginning and/or ending. Some ideas may seem out of order. Some problems with paragraphs.	There is no real beginning or ending. The ideas seem loosely strung together. Poor paragraph formatting.
<b>Voice &amp; tone</b>	The writing has a clear perspective, sophisticated style, and appropriate tone.	The style and tone are appropriate. The writer's perspective fades in and out.	The writer's perspective is obscure. The paper shows little awareness of audience and purpose.	The writing is flat, lacks a perspective, and uses an inappropriately formal or informal style and tone.
<b>Word choice</b>	The words used are descriptive but natural, varied and vivid.	The words used are correct, with a few attempts at vivid language.	The words used are ordinary. Some may sound forced or clichéd.	The same words are used repeatedly, some incorrectly.

Dimensions	4	3	2	1
<b>Sentence fluency</b>	Sentences are clear, complete, begin in different ways, and vary in length.	Mostly well-constructed sentences. Some variety in beginnings and length.	Many poorly constructed sentences. Little variety in beginnings or length.	Incomplete, run-on and awkward sentences make the paper hard to read.
<b>Conventions</b>	Spelling, punctuation, capitalization, and grammar are correct. Only minor edits are needed.	Spelling, punctuation and caps are usually correct. Some problems with grammar.	There are enough errors to make the writing hard to read.	The writing is difficult to understand because of errors.

From Andrade (2013)

It is preferable for teachers to design rubrics collegially as a group rather than as individuals. Taking advantage of how school teams already work together, as well as ensuring that the appropriate content expertise is represented in the group is a useful operating procedure for rubric development (Brookhart 2013). There is no rule of thumb for the frequency with which teachers should use rubrics. The use of a rubric depends on the purpose for which is being used (Brookhart 2013). For example, a rubric may be used at regular intervals during a writing assignment or once each week to assess oral reading. Given the time and effort to develop quality rubrics, it will be important to verify whether the learning goal or standard is best assessed by a performance task and a rubric, so that the investment in rubric development is worthwhile (Arter and Chappuis 2006).

Rubrics can improve student performance, as well as monitor it, by making teachers' expectations clear and by showing students how to meet these expectations. When teachers provide an evaluation of student work using a rubric, students should be clear about what they need to do to improve in the future. Rubrics can also help support student self- and peer assessment. (See section on Student Involvement for more information on self- and peer assessment). Rubrics are particularly useful for assessing oral language development, particularly for ELs. For example, rubrics can focus teachers' attention on particular discourse practices, grammatical structures, and

vocabulary as they observe and listen to their students' during collaborative discussions, oral presentations, and informal conversations. These observations can then guide instructional decision-making, including how teachers structure conversations and productive group work, how they model different uses of English, and how they ensure that students receive ample exposure to rich oral language, including from peers. The CA CCSS for ELA/Literacy and the CA ELD Standards provide useful information for designing rubrics to gauge students' progress in oral language (including vocabulary and presentations), collaborative discussions, writing, and other areas of the curriculum. Since the two sets of standards correspond, teachers can work together to create streamlined rubrics using both sets of standards, as well as the standards' companion appendices and documents, so as not to create multiple rubrics for evaluating the same tasks.

### ***Portfolios***

Student portfolios are another useful source of evidence for making judgments about student learning. A portfolio is a systematic collection of student work and related materials that tells the story of a student's activities, progress and achievement in a given subject area (Arter and Spandel 1992; Venn 2000). Portfolios can provide a progressive record of student growth or they can be used to demonstrate mastery of specific learning goals and contain only samples of a student's highest achievement (Venn 2000). They can be considered either medium- or long-cycle assessments, depending on the length of the period of learning they cover.

Whatever the purpose of the portfolio there should be sufficient samples related to specific learning goals that enable an evaluation of either growth or achievement (Chappuis, and others 2012). The specific learning goals should be aligned to the standards and the evidence that is included in the portfolio should reflect either students' progress toward meeting standards or achievement of specific standards.

Portfolios can contain a range of evidence: student learning goals; samples of written work; images of work samples (e.g., digital images of models or other representations); audio samples (e.g., student narratives; oral presentations or read-alouds), video files (student performances; ASL-signed presentations); student reflections; teacher observations; teacher-student conference notes; and documentation

of any other assessment results. Digital portfolios allow students to both assemble and publicly present their work. Assembling a portfolio should directly involve students in selecting the content as well as reflecting on why the contents were selected, what they represent, and what they show about the student's learning (Arter and Spandel 1992; Chappuis, and others. 2012).

Some questions teachers should keep in mind when using portfolios are

- How representative is the work included in the portfolio of what students can really do?
- Do the portfolio pieces represent coached work, independent work, or group work?
- Do the portfolio pieces represent student language and literacy progress across the content areas?
- How well do the portfolio items match standards?
- Are there clear criteria for judging the work and do the criteria represent the most relevant dimensions of student work products?
- Is there a method for ensuring that evaluation criteria are applied consistently and accurately? (Arter and Spandel 1992)

It is important to ensure that well-developed criteria are used to evaluate what portfolio items reveal about student achievement and the scoring process in place, for example, if the portfolio is scored by one or more raters and when scoring is done. It is also important to decide if the portfolio is to be rated as a whole (for example, a portfolio of student writing exemplifying achievement relative to specific standards), or if the portfolio is to be rated as individual samples. An additional consideration is the weighting of portfolio items. For example, will video performances of students' spoken language be weighted more or less than their written artifacts?

Portfolios have the added benefit of providing valuable information about student progress to parents, particularly the parents of ELs and other language-minority students who may not be completely familiar with U.S. schooling practices and systems. When portfolios are designed to *tell the story* of student growth during a particular time frame, this communicates to parents how their children are developing in a variety of areas valued in the standards and curriculum. This information can help parents to

support their students to continue developing in these areas at home, thereby creating an opportunity for collaboration between schools and families.

### **Student Involvement**

Whatever the assessment cycle, one goal of assessment is to promote a positive orientation to learning for students. Assessment, particularly when stakes are attached, creates a strong reason for learning. However, assessment can also impact the learner's willingness, desire and capacity to learn (Harlen and Deakin Crick 2002). For example, if passing the test becomes the reason for learning, then students run the risk of developing a performance orientation, rather than a learning and mastery orientation (Ames and Archer 1988; Dweck 1999, 2006). Students with a performance orientation tend to use passive rather than active learning strategies, they avoid learning challenges, and their learning tends to be shallow rather than deep (Crooks 1988; Harlen and James 1997).

While teachers can help students learn, only the students can actually do the learning. For this reason, successful achievement of standards will require students to develop a learning orientation evidenced by an interest in learning and meeting challenges, and a belief that effort, engagement in learning, and the development of learning strategies can lead to increased achievement.

If students are involved in the assessment process, they are more likely to develop a learning orientation than if they are solely passive recipients of test scores. They are also more likely to develop the skills of setting goals, managing the pursuit of those goals and self-monitoring, all important 21st Century skills (NRC 2012). Active student involvement in the assessment process is a vital element in the development of student self-direction in learning. Feedback is a crucial key to student involvement in assessment because it is a critical factor in the development of students' insight into their own learning and understanding (NRC 1999; OECD 2005).

### **Feedback**

Feedback provides an indication to students of what they have done well—the degree to which they have met the learning goals—and what they can do next to improve their learning (Bangert-Drowns, and others 1991). Importantly, feedback from teachers or peers should avoid focusing on the student and on making comparisons

with other students (Black and Wiliam 1998; Wiliam 2007). Instead, the focus of feedback should be on the task, the processes the student uses, and the student's self-regulation (Kluger and DeNisi 1996; Hattie and Timperley 2007). As Wiliam (2011) suggests, feedback should prompt for the learner a *cognitive* reaction (focused on active steps to achieve mastery) and not an *emotional* reaction (focused on anxiety and embarrassment).

Long- and medium-cycle assessments usually produce a score indicating the status of achievement. While the scores typically tell students *what* they have achieved, they do not tell them *how* or *why* they achieved what they did. The role of teacher feedback in relation to these types of assessment results is to assist students to understand where they were successful or not, and to set some goals with the students so they know where they need to improve and have some ideas of how to do so. This approach will mean that teachers need to spend time with students discussing assessment results and setting goals and strategies for improvement. Even when teachers have used a rubric and provided an evaluative score, students still need feedback about how to improve. Although potentially time consuming, the pay-off for students is that assessment is more transparent, and the students are more oriented to goals and feel more ownership in future learning.

When considering what kind of feedback to give their EL students on their developing English language use, teachers should focus first and foremost on effective communication and meaning-making. They should take note of language resources (e.g., vocabulary, grammatical structures, discourse moves) students are using, which is information teachers can use to plan further language learning opportunities, either in-the-moment or at a later time. Teachers should encourage EL students to take risks when using English. This is possible when teachers establish a safe and supportive environment where students are free to make *mistakes*—that are in fact normal developmental steps—in approximating complex academic uses of language as they learn increasingly advanced English. For example, a student might ask, “How fast the lava go?” If a teacher stops to correct the student's grammar (to tell the student they must use the word *does*, for example), the focus on meaning can be lost, and the student may be discouraged from taking further risks. Instead, teachers should think

carefully about when and how to provide feedback on particular aspects of students' language use, including grammatical structures, vocabulary, and register. For example, the teacher may at that moment simply acknowledge the student's question and recast the statement, thereby providing implicit feedback ("That's a great question! How fast does the lava flow? Let's read to find out."). The teacher could also write the question s/he recasts on a chart or document reader in order to provide a visual reinforcement of the oral modeling, and if a list of student questions has been generated, the class could even chorally read the questions together. In addition, the teacher could take note of what the student said and make plans to address the grammatical structure of questions more explicitly during designated ELD. These examples do not imply that explicit attention to English language development should not occur during content instruction. Rather, the way in which language development is addressed needs to be carefully considered in order to maximize student meaning making and risk taking. Overcorrection, particularly when it feels to the student like ridicule (for example, repeatedly requiring a student to repeat utterances in grammatically correct and complete sentences or chastising them for not using standard English pronunciation), is not effective feedback. It can take the focus away from content knowledge development and discourage EL (and non-EL) students from participating in conversations or inhibit their desire to write their ideas freely, which impedes their English language development.

**Snapshot 8.7 Student Involvement in Assessment in Grade Four**

Miss Nieto, a fourth grade teacher, has a discussion with each of her students about their reading scores from the interim assessment. In her meeting with Henry, she notes that the student has done well on the items related to using explicit details about the text and summarizing central ideas and is on track to meet the associated standards. She also discusses with Henry that his scores indicate that he is not as strong in using supporting evidence to justify or interpret how information is presented. Miss Nieto and Henry have a conversation about why he thinks he scored lower on some items. He tells her that he thinks he is getting the idea of using evidence for justification but he still thinks it is difficult for him. She suggests that this should be something he consciously focuses on improving between now and the next interim assessment, and she gives him some ideas that can support his learning.

Feedback is particularly salient in the context of formative assessment. Students can receive feedback in three ways: from their teachers, from peers, and through their own self-assessment. The purpose of the feedback is to close the gap between the student's current learning status and the lesson goals (Sadler, 1989). It is critically important that students be given opportunities to *use* the feedback, otherwise it does not serve the intended purpose.

### ***Teacher Feedback***

Three questions provide a frame for feedback to students in formative assessment:

1. Where am I going?
2. Where am I now?
3. Where to next?

To answer the first question, both teachers and students need to be clear about the goal or target of the learning and what a successful performance of learning will be.

Answering the second requires teachers and students to elicit and interpret evidence of learning. In other words, they need to decide where the students' learning currently stands in relation to the learning goal. Answering the third question guides the student to take next action steps toward meeting the learning goal. Feedback addresses both the second and the third questions. The teacher provides feedback that indicates to the student where he or she has been successful and provides a hint or cue of what to do next.

#### **Snapshot 8.8 Teacher Feedback in Grade One**

Kathleen, a first grader, is preparing to read aloud to her teacher. Before she begins, Mr. Silverstein reminds her to think about the reading strategies they have been working on. The text states: *Fish swim in the river.* Kathleen, reading very slowly, says: *Fish...swim...in...the...water. No. That's not water. It doesn't begin with 'w.'* R (says letter name) *r* (letter sound)... *i...v... River! Fish swim in the river.* Mr. Silverstein provides feedback after the student finishes reading the sentence: *You did a very good job of using your decoding strategies to read the text accurately. Let's keep on reading and while you are reading think about: is what you are reading making sense, and does what you are seeing match with what you are reading? Just like you did when you noticed that water could not be the right word. Water made sense, but the letters indicated a different, equally sensible word: river.*

### **Peer Feedback**

Peers are also sources of feedback for learning. Peer feedback has a number of advantages both for those students providing the feedback as well as those receiving it. It involves thinking about learning and can deepen students' understanding of their own learning. Research shows that the people providing the feedback benefit just as much as the recipient, because they are forced to internalize the learning goals and performance criteria in the context of someone else's work, which is less emotionally charged than their own (William 2006). The same three questions listed above apply to peer feedback. (See section on Teacher Feedback.) Without clarity about the goal and the performance criteria peers will find it difficult to provide useful feedback to each other. Peers need to assess the status of classmates' learning against the same success criteria they use to check their own learning. Additionally, providing constructive feedback is a skill students need to learn, so instruction will need to focus on this as well. It is worth remembering that learners who are adept at giving and receiving feedback to complete learning activities are acquiring important 21st Century skills (NRC 2012).

#### **Snapshot 8.9 Peer Feedback in Grade Three**

In a third-grade class students are focusing on Speaking and Listening standard 3.4, one of several that focus on *presentation of knowledge and ideas*. Their learning goal is to write an informative speech to present to the class about a topic of interest to them. The criteria they have to bear in mind when writing their speeches include the following:

- Introduce your topic in a way that engages your audience
- Put your ideas in a logical sequence
- Make an impact on your audience with your ending

Once the students have an initial draft, they exchange their papers with a partner. Then the students provide each other with feedback. One student's feedback to her partner is: *I liked how you started your speech with a question...that's a good way of getting your audience's attention. I think your ideas are logical. I think it would be a better impact at the end of your speech if you go back to your question and maybe finish with a sentence that tells how you answered the question.*

### ***Self-Assessment***

Teacher and peer feedback are externally provided. When students are involved in self-assessment they are generating *internal feedback*. Generating and acting on internal feedback is a form of metacognition and self-regulation. Metacognition is basically thinking about one's thinking, and self-regulation refers to the ability of learners to coordinate cognitive resources, emotions and actions in the service of meeting learning goals (Boekaerts 2006). In the realm of 21st Century learning, metacognition and self-regulation are important skills (NRC 2012). The most effective learners are self-regulating (Butler and Winne 1995; Pintrich 2000; Schunk and Zimmerman 2008). Additionally, training students in metacognition raises their performance (e.g., Lodico, and others 1983) and helps them generalize what they have learned to new situations (Hacker, Dunlosky, and Graesser 1998). Because of the importance of metacognition and self-regulation to successful learning, teachers need to pay attention to ensuring students develop these skills in the context of language and literacy learning.

Self-assessment can be developed from the early grades onwards (Perry, and others 2002; Puckett and Diffily 2004). For example, a first grade teacher provides her students with a graphic organizer with the headings: *date, book title, my goal today as a reader, pages read, how well did I meet my goals?* She asks her students to set goals for their independent reading time each day, and at the end of the session to think about how well they met the goals. During the week, when she has individual reading conferences with students, she reviews the self-assessment sheets and where students have not met their goal she asks them what the student did or needs to do to improve. Together, they set a strategy for the student to focus on. As well as providing the students with the opportunity for self-assessment, the teacher offers advice on strategies for improvement, which in turn become part of the students' internal repertoire of strategies that they can employ on subsequent occasions. In effect, they are developing the skills of self-regulation.

Self-assessment becomes more sophisticated as students gain more experience with the skill. For example, in a ninth grade science class where the teacher is integrating ELA and science standards, the students are involved in a short research project on distinct regions of the brain. As called for in the ELA writing standards for

literacy in science (WHST.9-10.6), they are to display their information “flexibly and dynamically.” Students in this class have time toward the end of every session to complete a reflection and planning log where they answer the following questions: *What was successful about your learning today? What difficulties or problems did you encounter? How did you manage those difficulties? Were you successful? If not, what plans do you have for dealing with them in the next lesson?* These logs serve as a means of self-assessment for students and support self-regulation because they have to think about strategies to solve difficulties. The logs are also sources of information for teachers about the progress students are making on their projects.

### **Assessment of ELD Progress**

Assessing ELD progress, particularly the development of *academic uses of English* in each discipline, is a responsibility shared by all educators in schools and districts where ELs are students. (See Chapter 11 for information on district and school leadership responsibilities for monitoring ELD progress.) District- and schoolwide assessment and professional learning *systems* are critical for ensuring EL students’ readiness for college, career, and civic life. However, it is *teachers* who ultimately must ensure that every day, each of their EL students has full access to the grade level curriculum and that they develop academic English in a timely—and even accelerated—manner. Teachers’ deep understandings of the CA CCSS for ELA/Literacy, the CA ELD Standards, and other content standards are foundational for effective assessment **for** and **of** learning because these standards guide instructional and assessment practices with ELs.

The CA ELD Standards provide outcome expectations at different English language proficiency levels (Emerging, Expanding, Bridging) so that teachers can differentiate their instruction according to individual EL students’ language learning needs on particular standards. Because the CA ELD Standards delineate proficiency levels which EL students are expected to progress through during the year (and in fact, they may progress through more than one level in a single school year), teachers must carefully attend to the ELD progress of their EL students on a frequent and ongoing basis. As described earlier in this chapter, this ongoing monitoring of student progress involves using short-cycle formative assessment (minute-by-minute, daily, weekly), as

well as medium-cycle assessment for formative purposes (monthly, end-of-unit, interim, benchmark, and other periodic time frames). Attending to the developing capacities and emerging or persistent needs of ELs is intertwined with the assessment approaches teachers employ for all students. However, because ELs are learning English as an additional language *at the same time* as they are learning content knowledge through English (and therefore have particular English language learning needs), teachers need to take additional steps in assessing their ELD progress and acting on evidence gathered from assessment. When thinking about short- and medium-cycle assessment of ELD progress, teachers should ask themselves the following questions:

- How do I determine what my EL students' levels of English language proficiency (ELP: Emerging, Expanding, Bridging) are on different CA ELD Standards?
- How can I use information about my students' ELP levels on different CA ELD Standards, as well as other relevant information, to design and provide targeted instruction that fosters language-rich learning opportunities?
- How often should I assess ELD progress? Which kinds of evidence-gathering approaches and tools are most appropriate for different purposes?
- How will I know if my EL students are making sufficient progress in developing English on a daily or weekly basis and over longer periods of time?
- How can I include my EL students in assessing their own ELD progress and support them to be conscious of and intentional in their English language learning?

Guidance and two examples for addressing some of these questions follows. This guidance focuses on how classroom teachers—including ELA teachers, teachers in other content areas, ELD teachers, and EL specialists who support content teachers—can use the CA ELD Standards to assess the ELD progress of each of their EL students. The CA ELD Standards support teachers' formative assessment practices by offering rich descriptions of what EL students can be expected to do at the end of each English language proficiency level (Emerging, Expanding, Bridging). These expectations can help teachers focus their formative assessment practices (assessment *for learning*) as their EL students use English while learning is taking place; gauge their EL students' developing capacities in English; and adjust instruction and learning

opportunities while learning is still occurring. The CA ELD Standards can also support teachers in their assessment of learning and to use this information for formative purposes. For example, the CA ELD Standards can be used to develop tools for observing language learning in context and evaluate language learning after instruction has occurred. The following guidance illustrates ways in which evidence of language learning gathered from careful observation, evaluation, and analysis can inform purposeful instruction that supports further development of ELs' academic uses of English. The guidance and examples offered here are intended to be used *in addition to—not instead of*—those provided in the rest of this chapter.

### **Assessing ELD Progress in Writing**

One way in which teachers can observe and respond to how their EL students are developing written language is by using a *language analysis framework for writing*, based on the CA ELD Standards and tightly aligned to teachers' learning goals and success criteria for writing. A language analysis framework allows teachers to observe and analyze student language in *linguistic terms*, with more specificity than is often found in ELA rubrics or other tools for evaluating writing. For example, feedback to students on writing, such as “interesting beginning, developed middle, and satisfying conclusion,” “could use more varied sentence patterns,” or “needs some colorful vocabulary”) may be sufficient for some students to use to improve their writing. However, such feedback may not be explicit enough for many students, including ELs, to act on (Fang and Wang 2011). Teachers providing this type of feedback may know intuitively what kind of writing they would like to see their students produce because it is related to their learning goals and success criteria, but without specific feedback on the *language resources* that constitute “varied sentence patterns” or “colorful vocabulary,” for example, their feedback on where to go next is elusive to ELs, and such language use remains a “hidden curriculum” (Christie 1999).

A language analysis framework for writing, drawing from the CA ELD Standards and other resources focused on language development, can help teachers provide a level of explicitness about the *specific language resources* that students can use in their academic writing to meet teachers' identified learning goals and success criteria in different disciplines. A language analysis framework provides *framing questions* that

students can ask themselves as they are writing and that students and teachers can ask themselves as they examine writing. For students, such guidance for composing and revising their own writing can support them to structure their texts more cohesively and use the types of grammatical structures and vocabulary that are expected. By explicitly focusing on language, expectations for writing become more transparent. Teachers can also use a language analysis framework to determine how well students are using particular language resources in a piece of writing they are currently composing in order to provide useful feedback to students and adjust instruction accordingly. It can also be used to gather evidence on whether and to what extent students are using these language resources over time. An example of a language analysis framework for writing in the upper elementary grades, developed using the CA ELD Standards and their English language proficiency descriptors (CDE 2014), as well as research on language development, is provided in Figure 8.7.

Figure 8.7. Language Analysis Framework for Writing

Language Analysis Framework for Writing				
Content Knowledge and Register	Text Organization and Structure	Grammatical Structures	Vocabulary	Spelling and Punctuation
<p>Is the overall meaning clear?</p> <p>Are the big ideas there, and are they accurate?</p> <p>Is the text type (e.g., opinion, narrative, explanation) appropriate for conveying the content knowledge?</p> <p>Does the register of the writing match the audience?</p>	<p>Is the purpose (e.g., entertaining, persuading, explaining) getting across?</p> <p>Is the overall text organization appropriate for the text type?</p> <p>Are text connectives used effectively to create cohesion?</p> <p>Are pronouns and other language resources used for referring the reader backward or forward?</p>	<p>Are the verb types and tenses appropriate for the text type?</p> <p>Are noun phrases expanded appropriately in order to enrich the meaning of ideas?</p> <p>Are sentences expanded with adverbials (e.g., adverbs, prepositional phrases) in order to provide details (e.g., time, manner, place, cause)?</p> <p>Are clauses combined and condensed appropriately to</p>	<p>Are general academic and domain-specific words used, and are they used accurately?</p> <p>Are a variety of words used (e.g., a range of words for “small”: little, tiny, miniscule, microscopic)?</p>	<p>Are words spelled correctly?</p> <p>Is punctuation used appropriately?</p>

		join ideas, show relationships between ideas, and create conciseness and precision?		
--	--	---	--	--

From Spycher and Linn-Nieves (2014), adapted from Derewianka (2011), Gibbons (2009), and Spycher (2007)

Teachers can use such a framework (adjusted appropriately for grade level/span), accompanied by their knowledge of their students (including students' ELP level on different CA ELD Standards), *for observing* what students are doing while writing and *for evaluating* students' writing products after they have produced them. For observing *in-the-moment* writing, having a framework for analyzing writing helps teachers focus on one or two areas to provide *just-in-time* scaffolding. Students can also use a language analysis framework or related tools, such as a *success criteria* document (Heritage 2014) addressing particular language areas, to evaluate and refine their own writing. Tools such as these support students to reflect on their work and ask themselves the same types of questions—either as they are writing or during a writing conference with peers or teachers—that teachers ask themselves when analyzing student writing. Using a language analysis framework for discussing the language students are using in their writing helps students to monitor their own progress in writing.

The following annotated writing sample (Figure 8.8) illustrates how a teacher used a language analysis framework to analyze student writing in order to determine next steps for instruction. The example was written independently by an EL student in the fifth grade after several days of instruction where students jointly constructed several short sections of a longer text on bats (see Vignette 5.3 in Chapter 5 for an example of the task, “text reconstruction”). The students used their shorter reconstructed texts, along with other texts they had used to learn about the topic, to construct their own texts on bats at the end of the week. The teacher’s analysis was used for formative assessment purposes and for discussing refinements with students and not for grading the student’s writing.

Figure 8.8. Student Annotated Writing Sample Using the CA ELD Standards

Susana's Text	Annotations
<p><b>Bats</b></p> <p>Bats are important because they eat mosquitos, insects, mice, frogs and other small animals that could become pest in your house. They are also important because they spread pollen and seeds and because of that more plants grow. If it weren't for bats we wouldn't have all the food that we have now.</p> <hr/> <p>There are over 1,200 species of bats. The largest bat is the flying fox and you could find it in Australia. 3 species of bats are bumble bee bat, fruit bat, and even vampire bats. Bat can be able to damage many plants. They don't suck blood like in horror movies. They drink it like little kittens. They are the only mammals that could fly.</p> <hr/> <p>Bats are in danger because people are scared of them. They are scared of them that they burn there homes. There are less bats know than they used to be. In Australia flying foxes are dieing of heat waves. In 15 years over 30,000 bats are dying because of U.S.</p> <p><b>Summary Notes and Next Steps:</b></p> <p>Discuss with Susana:</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Ordering of the three chunks, need for introduction that foregrounds the chunks, conclusion that sums them up</li> <li><input type="checkbox"/> Review whether information in each chunk fits there and if ideas in each chunk could be expanded more</li> <li><input type="checkbox"/> Show where clauses are combined to show relationships between them (e.g., using <i>because</i>), and ask her to see where she could do the same to combine other clauses</li> </ul> <p>Discuss with the class (based on patterns in other students' writing):</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> how register shifts when <i>you</i>, <i>we</i>, <i>us</i> are used</li> <li><input type="checkbox"/> how connecting and condensing ideas (clause combining or other ways) creates relationships between ideas and reduces repetition (maybe a mini-lesson with examples from student writing we revise together)</li> <li><input type="checkbox"/> how to use text connectives (maybe revise a piece of writing together and add in text connectives where</li> </ul>	<p><b>Content and register:</b></p> <ul style="list-style-type: none"> <li>big ideas and lots of informative details provided, mostly accurate information</li> <li>some information needs more clarity (bats aren't in danger just because people are scared of them)</li> <li><i>you</i>, <i>we</i>, <i>us</i> is used (less formal register)</li> </ul> <p><b>Text structure and organization:</b></p> <ul style="list-style-type: none"> <li>organized logically into three chunks (<i>why bats are important</i>, <i>species of bats</i>, <i>why bats are in danger</i>)</li> <li>some information doesn't seem to fit in the chunks (<i>bats damaging plants</i>)</li> <li>missing an introduction and conclusion, order may not be logical</li> <li>pronoun reference: <i>because of that</i> used accurately to condense and link to previous sentence (cohesion)</li> <li>could use more text connectives (cohesion)</li> </ul> <p><b>Grammatical Structures</b></p> <ul style="list-style-type: none"> <li>some appropriate clause combining to link ideas and show relationships</li> <li>some clause combining needs work (<i>They are scared ... that they burn ...</i>) and more could be used</li> <li>phrases could be expanded to include more details about where, when, etc.</li> </ul> <p><b>Vocabulary:</b></p> <ul style="list-style-type: none"> <li>domain-specific (mammals, species, pollen) and general academic (spread, damage) vocabulary used accurately</li> </ul> <p><b>Spelling and punctuation:</b></p> <ul style="list-style-type: none"> <li>mostly accurate, with some approximations (<i>mamles</i>, <i>dieing</i>)</li> </ul>

Using a language analysis framework for writing also helps focus discussions about writing. For example, in a writing conference during which Susana has an opportunity to discuss her writing with her teacher, Susana's teacher might open the conversation by asking Susana to identify areas where she felt her writing needed refinement. Susana's teacher might continue the conversation by acknowledging specific areas where Susana's writing was strong (e.g., "I see that you are providing lots of great content information about bats and that you're organizing the information in a way that helps the reader follow your ideas."). She might ask probing questions to prompt Susana to notice areas for refinement or explicitly point out specific places in the text that need refinement, areas that are important for Susana to develop as a writer but that Susana herself may not be aware of. She would use the CA CCSS for ELA/Literacy for grade five to frame her learning goals for the conference and the CA ELD Standards to help her to provide targeted support, based on Susana's ELP level on particular CA ELD Standards. Some of the questions she might ask to prompt Susana's thinking and extend her use and understanding of English are the following:

- How could you orient the reader to what your paper is about? How could you let them know in advance about the categories you've chosen to include?
- Does all of this information belong together in this section?
- How could you expand this idea to add more detail?
- How could you combine these ideas to show the relationship between them?
- Is there another word or phrase that would help you get your meaning across in a more precise way?

After examining evidence of student writing, teachers can determine whether and in what ways students have progressed, and what next steps are needed in instruction to support further language learning. For example, if a group of EL students at the Emerging level of ELP are not yet using pronouns to refer to information that has already been presented in a text, their teachers might model how to do this, provide many opportunities for them to apply this new language resource to their own writing, and continue to draw their attention to pronoun reference until the students have internalized this understanding. If EL students at the Expanding level are already using pronoun reference but not yet using more sophisticated cohesive language resources,

such as the use of demonstratives (e.g., this, that) or nominalization to (e.g., *the result of environmental degradation ...*), their teachers can show them examples of these language resources used in the texts they are reading, have students analyze texts for such language resources, and provide multiple opportunities for students to apply this awareness of how English works to their own writing. Teachers can monitor how well students *take up* these language resources in their writing over time and provide targeted feedback to the whole class, small groups, or individual students so that they can continue to progress in their English language development.

### **Assessing ELD Progress in Oral Language**

Oral language use is a critical component of English language development, and observing how students are developing language skills, abilities, and awareness needed for collaborative conversations and other oral language tasks, such as oral presentations, is essential. In carefully planning collaborative learning opportunities and intentionally observing their EL students as they engage in these tasks so that they can provide just-in-time scaffolding to stretch students' oral language to higher levels of proficiency, teachers engage in assessment for learning. These formative assessment practices, which should remain the top priority during classroom instruction, are complemented by more formal evidence-gathering strategies and tools for observing and documenting progress in English oral language development.

For example, in grade seven, students are expected to engage in small group discussions about complex texts. One of their conversations might revolve around an informational science text they are currently reading. As the students discuss their ideas about the text and extend their thinking about the content, the focus of teachers' observations is primarily meaning making. In other words, are the students showing that they are understanding the content of the text, making appropriate inferences based on textual evidence and background knowledge, using relevant examples, and extending their own understandings of the text by asking their peers questions and answering questions posed to them?

Yet teachers can also observe how their EL students are using English to convey their ideas and engage in academic conversation in the context of authentic, meaningful interaction about complex texts and topics. Clearly, teachers would also be observing

their non-EL students' academic language development during these meaningful interactions with texts, tasks, and others. However, the CA ELD Standards specifically help teachers to determine, by English language proficiency level (Emerging, Expanding, Bridging), the types of language resources their EL students should be able to use in collaborative conversations. This could form the basis for evidence-gathering strategies and tools that help focus observations and determine next steps for supporting their oral language development.

Such strategies and tools can be used to focus attention on specific language uses that teachers and students, as appropriate based on age, determine are areas of growth. Such observation tools help teachers to notice how their students are progressing in their capacity to engage in collaborative conversations. As teachers develop deeper understandings of the CA ELD Standards, they begin to notice how their EL students are using English in the context of specific CA ELD standards. They also become more skilled at identifying where on the ELD continuum their students are and where they need to go next in their academic uses of language. These observations help to anchor teachers' monitoring of ELD progress.

Observation tools should be used strategically and purposefully. For example, teachers might use a formal observation tool monthly or quarterly. For some students, the tool may be used more frequently (for newcomer ELs at the early Emerging level, for example) or less frequently (for students at the late Bridging level, for example) because the tool is intended to complement the ongoing observations teachers make every day. In addition, oral language observation tools support teachers to make informed instructional choices based on the evidence gathered using the observation tool. They are not intended to be used for grading students. Figure 8.9 provides an example of one such observation tool that a team of seventh grade teachers might use to monitor how their EL students at different places along the ELD continuum use English in collaborative conversations.

Figure 8.9. Grade Seven Collaborative Conversations Observation Notes

Collaborative Conversations Observation Notes			
<b>English Language Development Level Continuum</b> →---- Emerging -----→----Expanding -----→----- Bridging -----→			Students said ... (note students' comments and names)
<b>CA ELD Standards in Focus:</b>			
<b>Exchanging Ideas Respectfully (ELD.PI.7.1)</b>			
Engage in conversational exchanges and express ideas on familiar topics by asking and answering yes-no and wh- questions and responding using simple phrases.	Contribute to class, group, and partner discussions by following turn-taking rules, asking relevant questions, affirming others, adding relevant information, and paraphrasing key ideas.	Contribute to class, group, and partner discussions by following turn-taking rules, asking relevant questions, affirming others, adding relevant information and evidence, paraphrasing key ideas, building on responses, and providing useful feedback.	
<b>Supporting Opinions and Persuading Others (ELD.PI.7.3)</b>			
Negotiate with or persuade others in conversations (e.g., to gain and hold the floor or ask for clarification) using learned phrases (e.g., <i>I think. . . , Would you please repeat that?</i> ) and open responses.	Negotiate with or persuade others in conversations (e.g., to provide counter-arguments) using learned phrases ( <i>I agree with X, but . . .</i> ), and open responses.	Negotiate with or persuade others in conversations using appropriate register (e.g., to acknowledge new information) using a variety of learned phrases, indirect reported speech (e.g., <i>I heard you say X, and I haven't thought about that before</i> ), and open responses.	
<b>Connecting Ideas (ELD.PII.7.6)</b>			
Combine clauses in a few basic ways to make connections between and join ideas (e.g., creating compound sentences using <i>and, but, so</i> ; creating complex sentences using <i>because</i> ).	Combine clauses in an increasing variety of ways (e.g., creating compound and complex sentences) to make connections between and join ideas, for example, to express a reason (e.g., <i>He stayed at home on Sunday in order to study for Monday's exam</i> ) or to make a concession (e.g., <i>She studied all night even though she wasn't feeling well</i> ).	Combine clauses in a wide variety of ways (e.g., creating compound, complex, and compound-complex sentences) to make connections between and join ideas, for example, to show the relationship between multiple events or ideas (e.g., <i>After eating lunch, the students worked in groups while their teacher walked around the room.</i> ) or to evaluate an argument (e.g., <i>The author claims X, although there is a lack of evidence to support this claim.</i> ).	
<b>Quick Observation Analysis:</b>			
<b>Next steps to take:</b>			

The tool provided above is used to complement the more informal minute-by-minute observations teachers are already making of their students during collaborative conversations. Of course, care should be taken in implementing such tools. For example, attempting to observe too many standards at once or using the tool too often can be frustrating and counter-productive. Teachers will need support and the flexibility to use such tools in ways that best inform their instructional practice.

The approaches and tools for assessing ELD progress provided in the preceding pages are simply illustrative of how teachers can attend to their EL students' progress in developing English as an additional language. These tools are not meant to be prescriptive. Teachers should develop and employ assessment approaches that support the learning goals they have for all students and strategically select additional approaches (when needed) that will help them to ensure their EL students are advancing along the ELD continuum in a timely manner.

Another critical area of monitoring ELD progress for schools and districts is the appropriate use of large-scale summative assessment evidence, such as the CELDT. As delineated in Figure 8.3 earlier in this chapter, such summative assessments are not intended for planning daily instruction. Rather, the evidence from large-scale summative assessments related to ELD helps schools and districts evaluate and adjust the design of instructional programs provided to ELs and measure ELs' progress in learning English *from year-to-year*. Systematic monitoring of multi-year ELD progress should determine if EL students are progressing in their English language development within appropriate time frames and employ clearly defined protocols for action if they are not.

For example, a school leadership team would conduct a systematic and careful analysis of year-to-year ELD progress, based on the current and several prior years of summative assessment results (in concert with other measures of student achievement), to identify EL students who

- Appear to be ready to reclassify as English proficient;
- Are progressing in their English language development at an appropriate rate; or
- Have stalled in their ELD progress.

The team would carefully identify possible reasons to explain progress or the lack thereof. Using the results of their analyses, the team would determine specific and

timely next steps for instructing individual students, as well as appropriate adjustments and additions to program design, professional learning, and the school or district's comprehensive assessment system. In addition, the team would ensure that there is an accountability system in place to measure the efficacy of these adjustments and additions. Additional guidance on reclassification is provided in Chapter 11.

### **Assessment for Intervention**

Screening, diagnostic, and progress-monitoring assessments are discussed in this section. Screening assessments identify students who may have difficulties, diagnostic assessments give specific information about the difficulties, and progress-monitoring assessments provide feedback on whether planned interventions to address the difficulties are working. These assessments can operate in short or medium cycles.

#### **Universal Screening (Medium Cycle)**

Universal screening is a critical first step in identifying students who are at risk of experiencing reading difficulties and who may need more instruction. Universal screening consists of brief assessments focused on target skills (e.g., phonological awareness) that are highly predictive of future outcomes (Jenkins 2003).

An expert panel convened by the U.S. Department of Education's Institute of Education Sciences recommended that screening should take place at the beginning of each school year in kindergarten through grade two, and a second screening mid-year for kindergarten and grade one (Institute of Education Sciences [IES] 2009).

Because of students' development, the panel also recommended target areas for early screening. Kindergarten screening batteries should include measures assessing letter knowledge, phonemic awareness, and expressive and receptive vocabulary. As children move into grade one, screening batteries should include measures assessing phonemic awareness, decoding, word identification, and text reading. By the second semester of grade one the decoding, word identification, and text reading<sup>3</sup> should include speed as an outcome. Grade two batteries should include measures involving word reading and passage reading. For a reasonably accurate identification of students,

---

<sup>3</sup> As noted earlier, fluency rates do not apply to students who are deaf and hard of hearing who use American Sign Language as they are actually translating from one language to another when they storysign.

the panel also recommended the use of two screening measures at each juncture. When schools or districts are selecting screening measures they should carefully examine the technical information available from the publisher's manual (IES 2009).

### **Diagnostic Assessment (Medium Cycle)**

While the purpose of diagnostic assessments is to improve student learning, they should not be confused with short-cycle formative assessment. Formative assessment is used to guide ongoing decisions about student learning, whereas diagnostic assessment is used to identify areas where intervention may be needed to improve student learning. (Carnegie Council on Advancing Adolescent Literacy 2010).

Poor performance might reflect any one of a number of problems including, but not limited to, struggles with language and literacy. For example, if students are having difficulty understanding grade-level text, they may have short-term memory issues, may not read fluently enough to focus their attention on meaning making, or may not be making connections across phrases and sentences in the text. Diagnostic assessment is the means by which to identify the precise source(s) of the student's difficulty so that an appropriate intervention can be planned. Timely identification of students' difficulties is essential to ensuring the right intervention is made so students can progress.

Great care should be taken when approaching diagnostic assessments in English for ELs and students who are deaf. For example, an EL at the Emerging level of English language proficiency or a student who is deaf may appear to struggle with reading comprehension when reading a complex text in English. However, it could be that the student has not had sufficient opportunity to build up the language resources in English (including vocabulary and grammatical structures) or background knowledge needed to apply reading comprehension strategies. With appropriately adjusted instructional support, the students may demonstrate comprehension. Diagnostic assessments administered in English to ELs and students who are deaf need to be interpreted carefully. Teachers should consider possible linguistic and cultural biases of assessments (see section on Technical Quality in this chapter), use multiple types of assessments (including, where appropriate, assessments given in the primary language) to gain a comprehensive portrait of students' learning needs, and compare the student to her or his peers who are ELs or who are deaf and not just to native

English speakers.

According to the *Standards for Educational and Psychological Testing* (American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) 1999), any test that uses language is in part a test of language. Therefore, for ELs and students who are deaf who use American Sign Language, every test written in English—regardless of the content area—is partially a test of their English language proficiency and may not adequately assess their content area knowledge and skills (Abedi 2002). For this reason, it may be beneficial to assess them in their primary language in order to gain a more complete picture of their strengths and needs. However, it is important to bear in mind that it may not be appropriate to use content assessments in the primary language with every EL student. For example, students who are literate or are receiving formal instruction in their native language in this content area and who are at lower ELD levels are more likely to benefit from a content assessment in the primary language than those who are not (Pennock-Roman and Rivera, 2011; Bowles and Stansfield 2008; Stansfield and Bowles 2006). In the same vein, evaluating emerging bilinguals' writing by looking at their Spanish writing side by side with their English writing can help teachers see how the languages reinforce each other, and provide a bigger picture view of the students' developing biliteracy (Soltero-Gonzalez, Escamilla, and Hopewell 2012).

A range of assessments is available for diagnosing the source of a student's difficulties and it will be important to ensure the appropriateness of these assessments for diagnostic purposes. (See the section on Technical Quality in this chapter). Administering and interpreting some diagnostic assessments requires special training and licensure so when selecting diagnostic assessments it will be important to determine if the school has access to professionals who are qualified to administer them. Teachers can benefit from working closely with reading specialists who have the necessary specialized knowledge to interpret diagnostic data and provide guidance regarding specific interventions (International Reading Association 2000). It is advantageous for all available professionals (e.g., teacher, reading specialist and school psychologist) to work together in diagnosing a student's problem and planning

appropriate interventions (Joseph 2002).

### **Progress Monitoring (Short or Medium Cycle)**

Progress monitoring (sometimes referred to as curriculum-based measurement or curriculum-based assessment) is the practice of assessing students' academic performance on a regular basis for three purposes: 1) to determine whether students are profiting appropriately from the instructional program, including the curriculum; 2) to create more effective programs for those students who are not benefitting; and 3) to estimate rates of student improvement (National Research Center on Learning Disabilities 2006). To implement progress monitoring, a student's current level of performance is determined and goals are established for learning that will take place over a specific period. The student's academic performance is assessed on a regular basis (see IES 2009 recommendations above) and progress toward meeting the goal is determined by comparing the actual and expected rates of learning.

In addition to the general screening measures described above, a system of progress monitoring is recommended in response to intervention (RTI) programs (IES 2009). Based on available evidence, the panel convened by IES recommended that progress-monitoring assessments be administered to Tier<sup>4</sup> 2 students at least once each month. For those students who are not making sufficient progress, a Tier 3 intensive intervention will need to be planned. Progress-monitoring assessments should be used in Tier 3 to determine the effectiveness of the intervention (IES 2009).

National Association of State Directors of Special Education (NASDSE) identified nine essential characteristics for progress monitoring to be useful in an RTI context. Recommendations include that progress monitoring should assess marker variables that have been demonstrated to lead to the ultimate instructional target, be sensitive to small increments of growth over time, be administered repeatedly using multiple forms, be administered efficiently over short periods, and result in data that can be summarized in teacher-friendly data displays (NASDSE 2005, 25–26).

If teachers, schools, or districts wish to adopt progress-monitoring assessments, careful attention will need to be paid to the technical quality of any proposed

---

<sup>4</sup> See Chapter 9 for a discussion of tiers.

assessments to ensure they are appropriate for the intended purpose. (See the section on Technical Quality in this chapter.)

### **Mandated California Assessments**

On October 2, 2013, AB 484 established the California Assessment of Student Performance and Progress (CAASPP) assessment system, which replaces the Standardized Testing and Reporting (STAR) program. The primary purpose of the CAASPP system is to assist teachers, administrators, and students and their parents by promoting high-quality teaching and learning through the use of a variety of assessment approaches and item types.

Beginning in the 2014-2015 school year, student performance in grades three through eight and in grade eleven will be assessed by annual assessments developed by the Smarter Balanced Assessment Consortium and administered in accordance with CAASPP regulations, %CCR Section 855(b)(1) and (2). See Figure 8.10. This new state law exempts ELs from taking the ELA portion of the SBAC assessment if they have been enrolled in a U.S. school for less than 12 months.

To ensure the assessments address the full range and depth of the CA CCSS for ELA/Literacy, and the breadth of achievement levels, Smarter Balanced assessments combine item types, including selected response (multiple-choice items with one or multiple correct responses and two-part items) and constructed response (students write a short text or long essay in response to a prompt). For example, for the third grade reading standard, *determine the main idea of a text; recount the key details and explain how they support the main idea* (RI.3.2), selected-response items could be used to assess *determine the main idea of a text; recount the key details*, while a constructed-response item could be used to assess *explain how they support the main idea*. A computer-adaptive assessment, item response types will also include matching tables, fill-in tables, select or order text or graphics, and drag and drop.

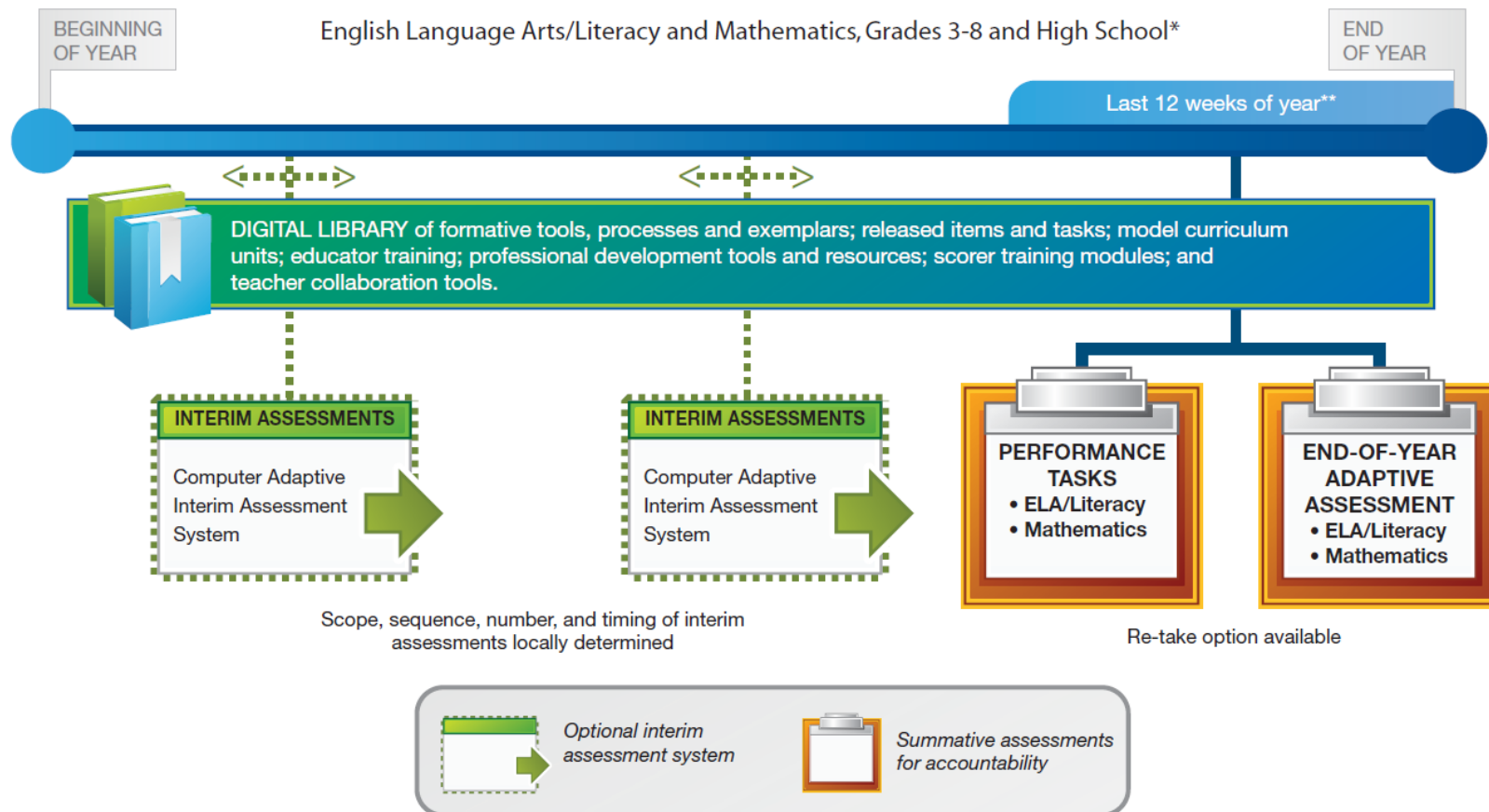
For results to be timely and useful, achievement of students in kindergarten through second grade, and reading standards for foundational skills for kindergarten through grade five that are critical to every student's success in reading, are typically assessed through locally determined assessment. The foundational skills are assessed intensively at kindergarten through grade two and then strategically at grade levels

above grade two. In selecting appropriate assessments for the purpose of assessing kindergarten through second grade students' achievement relative to standards, and assessing foundational skills, it will be important to refer to the section on the technical quality of assessments in this chapter to ensure that the assessments used are appropriate for their intended purposes.

Optional interim assessments are also available to be administered at locally determined intervals. The interim assessments are reported on the same scale as the year-end assessments and permit teachers to assess either clusters of standards (referred to as Block Assessment) or the full range of the CA CCSS ELA/Literacy Standards (referred to as Comprehensive Assessment). In addition, Smarter Balanced has a digital library of formative practices and tools for teachers' use. These tools include model units and lessons with embedded formative assessment strategies for teacher use.

The Smarter Balanced end-of-year and eventually the interim assessments comprise computer adaptive tests and performance tasks, which are described in more detail below.

Figurer 8.10. Smarter Balanced Assessment Consortium System



\* Summative and interim assessments for grades 3 – 8 and 11, with additional supporting assessments available for grades 9, 10, and 12.

\*\* Grades 3-8: Testing shall not begin until at least sixty-six percent (66%) of a school's annual instructional days have been completed, AND testing may continue up to and including the last day of school. Grade 11: Testing shall not begin until at least eighty percent (80%) of a school's annual instructional days have been completed, AND testing may continue up to and including the last day of school.

Source: Center for K-12 Assessment & Performance Management at ETS (2014) [permission to be sought]

### **Computer Adaptive Tests**

Computer-adaptive tests (CAT) tailor an assessment to individual students by presenting items based on a student's performance or responses to previous items in the test (Smarter Balanced 2013a). The Smarter Balanced summative assessments are being developed for use with CAT technology known as computer adaptive testing. The CAT assessment "engine" begins by delivering a short series of moderately difficult grade-level test items to the student, and then, depending on the student's initial performance, delivers items that are either more or less difficult. This process continues until the student's level of proficiency is determined (Smarter Balanced 2013a). For example, if a student has performed well on prior items, then more difficult items will be given thereafter, but if a student has performed poorly on prior items, then easier items are presented to the student. By matching the difficulties of new items more closely with a student's demonstrated level of performance, fewer items are needed. Some of the competencies assessed by CAT items include students' ability to use evidence to support their analyses (i.e., claims, conclusions, inferences) from reading different levels of text and their ability to edit and revise writing samples of different levels of complexity.

Because the test is taken on the computer, it is critical that students have developed the necessary technology skills, such as keyboarding, manipulating a mouse, and using pull-down menus as well as ample experience with the devices they will use during the summative assessment.

### **Performance Tasks**

Performance tasks provide opportunities for students to demonstrate learning in ways that "emulate the context or conditions in which the intended knowledge and skills are actually applied" (AERA, APA, and NCME 1999, p. 137). They can take the form of demonstrations, oral performances<sup>5</sup>, investigations and written products (Lane 2013). Performance assessments provide better possibilities to measure complex skills and communication, important competencies and disciplinary knowledge needed in today's society (Palm 2008) and important learning goals that cannot be easily assessed with

---

<sup>5</sup> The term *oral language* refers to signed language for students who are deaf and hard of hearing who use American Sign Language as their primary language.

other formats (Resnick and Resnick 1992).

The Smarter Balanced performance tasks, some of which are lengthy and will take considerable time to complete, emphasize deep knowledge of core concepts and ideas, analysis, synthesis, communication and critical thinking. For example, to assess the writing standards across all grade levels, full compositions, involving planning and revision are assessed with performance tasks. Similarly, performance tasks are used to assess grade 6-12 reading and writing standards for literacy in history/social studies, science and technical subjects. For instance, short research projects that involve applying research and inquiry as well as a demonstration of many 21st Century skills to produce a range of products (e.g., script for a presentation, PowerPoint, public service announcement) are assessed with end-of-year performance tasks. Other constructed-response tasks include asking students to respond to a question about a passage they have read and use details from the text to support their answer, to write an ending to story by adding details to tell what happens next, revising a paragraph by adding details to support an argument, and highlighting parts of a text that provide evidence to support a core idea of the text.

### **Assessments for Students with Significant Cognitive Disabilities**

The Common Core State Standards are for every student, including students with significant cognitive disabilities. All students with disabilities will take the new assessments, with the exception of students who cannot achieve at or near grade level as identified by the members of the IEP team. These students present the most significant cognitive disabilities and make up approximately one percent of the population. They will require substantial supports under an alternative assessment. These supports will allow them to have meaningful access to certain standards and assessment experiences that are appropriate to the students' academic and functional needs. On October 1, 2012, California joined the National Center and State Collaborative (NCSC) Consortia. The NCSC is developing professional development modules and curriculum/instruction resources, creating alternate achievement standards and developing a multi-state comprehensive assessment system for students with significant cognitive disabilities. The long-term goal is to ensure that students with significant cognitive disabilities achieve increasingly higher academic outcomes and

leave high school ready for post-secondary options. The curriculum, instructional materials, and assessments targeted for students with significant cognitive disabilities are currently being developed. The NCSC is a standards-aligned assessment that is targeted to replace the previous alternate performance-based assessment known as the California Alternate Performance Assessment (CAPA). For more information, contact the California Department of Education Common Core Resources for Special Education website <http://www.cde.ca.gov/sp/se/cc/>.

### **Biliteracy Assessment**

When instruction is provided in English and in an additional language in alternative bilingual or dual language programs, classroom assessment for academic and language development progress in both languages is necessary. Such assessments should be designed according to the same principles and recommendations articulated throughout this framework and throughout this chapter for both ELs and for students whose primary language is English. Frequently and closely monitoring students' progress, assessing in both languages used for instruction, and interpreting assessment results in accordance with the research on effective bilingual education practices helps to ensure that students make steady and consistent progress toward full biliteracy and academic achievement in both languages. (Note that EL students who have been enrolled in a U.S. school for less than 12 months do not take the ELA portion of the SBAC summative assessment.)

### **English Language Proficiency Assessments**

The English Language Proficiency Assessments for California (ELPAC), which will be aligned to the CA ELD Standards adopted in 2012, is being developed to replace the California English Language Development Test (CELDT). The CELDT will be administered as usual until the ELPAC is fully operational.

The ELPAC will consist of two separate assessments: an initial assessment and a summative assessment. Below is a summary description and figure of the identification and assessment process for ELs:

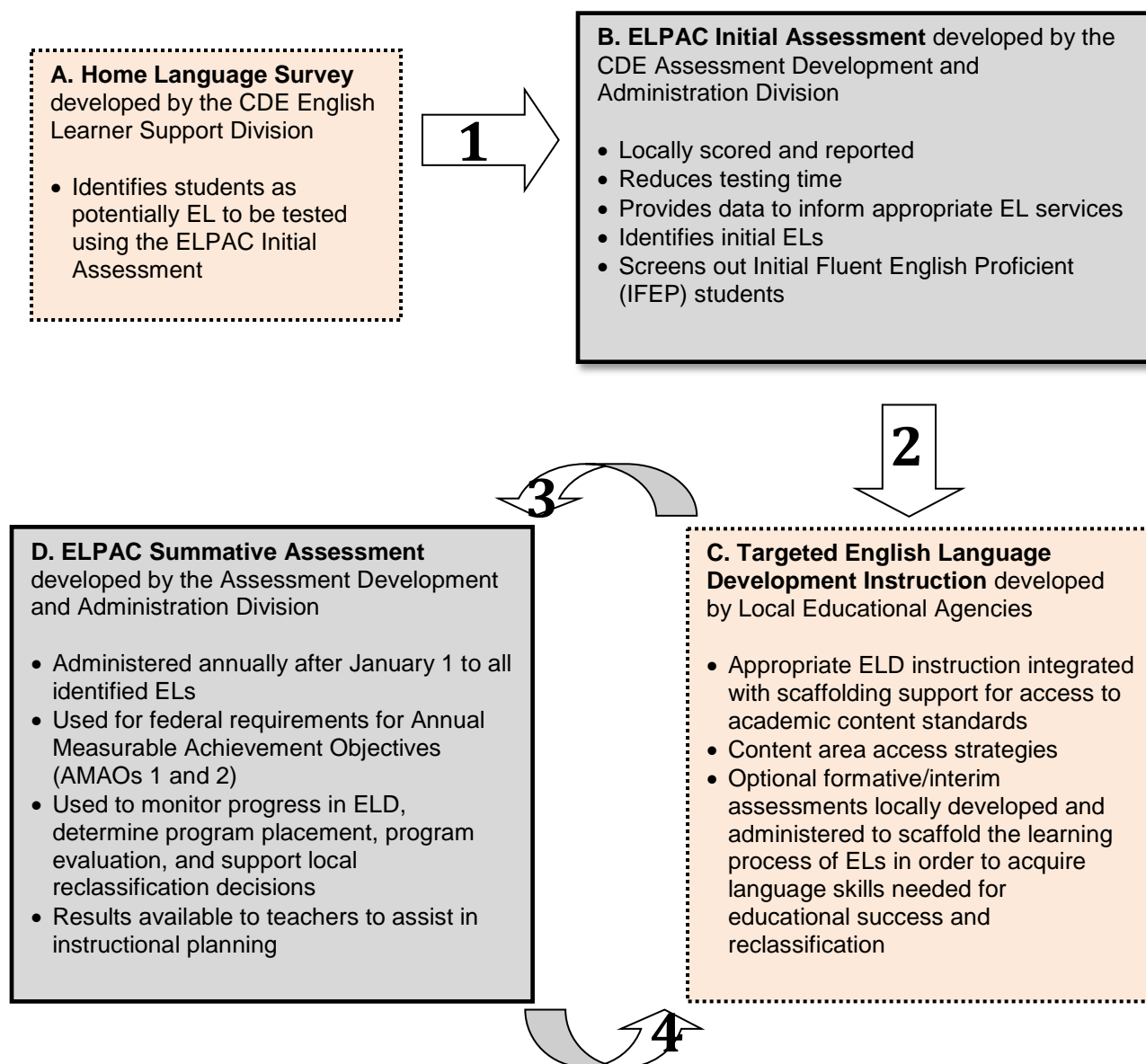
- **Home Language Survey (HLS):** School districts will continue to employ an HLS as the first step in identifying students whose primary language is not English. The HLS indicates if a student speaks a language other than English at home

sometimes or all of the time. This helps to determine which students are potentially EL and therefore require the ELPAC Initial to confirm EL classification.

- **ELPAC Initial Assessment (Initial):** The Initial will be used by school districts to determine whether or not a student is an EL. Scoring of the Initial will be done at the local level by trained ELPAC examiners, resulting in a quicker turnaround of test results and timelier determination of EL classification and placement of students in appropriate instructional programs.
- **ELPAC Summative Assessment (Summative):** Trained ELPAC examiners will annually administer Summative to all identified ELs during a four-month period after January 1 determined by the State Superintendent with the approval of the State Board of Education. The results will be used to determine ELs' annual progress toward learning English for federal accountability purposes. The results may also be used by school districts to evaluate the effectiveness of their ELD programs, curricular resources, and instruction.

The ELPAC conceptual model (Figure 8.11 below) highlights the process for using an HLS, the Initial, and the annual Summative. Boxes A and C have dotted borders indicating activities at the local level, and Boxes B and D have solid borders because they describe integral components of the State assessment system.

Figure 8.11. ELPAC Conceptual Model



Source: California Department of Education

### Technical Quality of Assessments

When considering the use of the Smarter Balanced Assessment System and other assessments to support student achievement of the CA CCSS for ELA/Literacy and the CA ELD Standards, it is important to keep in mind the purpose for which a given assessment is intended. If an assessment does not permit proper inferences and

provide accurate information for the specific decision-making purpose, its use may constitute misuse (Herman, Aschbacher, and Winters 1992).

This section elaborates the idea of the intended purpose of assessment. It will be particularly important to refer to this section when selecting assessments other than the Smarter Balanced assessments or the CELDT (and the ELPAC when it replaces the CELDT) whose technical quality has already been established through rigorous studies.

### **Elements of Technical Quality**

The idea of the *technical quality* of assessment refers the accuracy of information yielded by assessments and the appropriateness of the assessments for their intended purposes. There are three key elements related to the technical quality of assessments: validity, reliability, and freedom from bias (AERA, APA, and NCME 1999). Each element is described here, and Figure 8.12, summarizing the key points for each, is included at the end of this section.

#### ***Validity***

Validity is the overarching concept that defines quality in educational measurement. It is the extent to which an assessment permits appropriate inferences about student learning and contributes to the adequacy and appropriateness of using assessment results for specific decision-making purposes (Herman, Heritage, and Goldschmidt 2011). No assessment is valid for all purposes. While people often refer to the validity of a test, it is more correct to refer to the validity of the inferences or interpretations that can be made from the results of a test. Validity is basically a matter of degree; based on its purpose, an assessment can have high, moderate or low validity. For example, a diagnostic reading test might have a high degree of validity for identifying the type of decoding problems a student is having, a moderate degree for diagnosing comprehension problems, a low degree for identifying vocabulary knowledge difficulties and no validity for diagnosing writing conventions difficulties. Similarly, the annual end of sixth grade assessments will have a high degree of validity for assessing achievement of standards for those students, but no validity for assessing the incoming group of sixth graders' achievement.

For an assessment to be valid for the intended purpose, there should be evidence that it does, in fact, assess what it purports to assess. Test publisher manuals

should include information about the types of validity evidence that have been collected to support the intended uses specified for the assessment.

### ***Reliability***

Reliability refers how consistently an assessment measures what it is intended to measure (Linn and Miller 2005). If an assessment is reliable, the results should be replicable. For instance, a change in the time of administration, day and time of scoring, who scores the assessment, and any changes in the sample of assessment items should not create inconsistencies in results.

Reliability is important because it is a necessary adjunct of assessment validity (Linn and Miller 2005). If assessment results are not consistent, then it is reasonable to conclude that the results do not accurately measure what the assessment is purported to measure. A general rule of thumb for reliability is that the more items on an assessment the higher the reliability. Reliability is assessed primarily with statistical indices. Publishers' manuals should provide information about the reliability evidence for an assessment and the relevant statistical indices.

A variety of factors can influence the reliability of an assessment. For example, if a test is administered in an extremely hot or noisy room, students may not be able to complete the test to the best of their ability. If students are asked to provide an oral presentation when the instructions or expectations have not been made clear, this affects the reliability of the performance assessment. A number of other factors, including students' health, level of stress, and motivation can affect the reliability of an assessment. Teachers should use their judgment in interpreting assessment results when they suspect students are not able to perform to the best of their abilities. It is equally important for teachers to understand that a test or performance assessment may be reliable but not valid. For example, a student may consistently do well on an assessment, but the assessment may not be measuring what it claims to measure.

### ***Freedom from Bias***

Bias can occur in test design or the way results are interpreted and used. Bias systematically disadvantages a student or group of students so that the student(s) are unable to accurately show what he or she knows and can do with respect to the content of the assessment. As a result, the assessment results may underestimate the student's

achievement or reflect abilities that are not related to the assessment's content (Abedi and Lord 2001). Bias arises from tests that favor students of a particular gender, ethnicity, cultural background, geographic location, disability and primary language. An assessment that is free from bias will produce the same scores for students of the same attainment level, irrespective of their demographic subgroup.

Popham (1995) identifies two forms of bias, offensiveness and unfair penalization. Offensiveness occurs when the content of an assessment offends, upsets, or distresses particular subgroups, thus negatively influencing the test performance of these students. Items that present stereotypes of girls, boys, or particular cultures, or that portray certain groups as inferior, could adversely affect certain students' performance.

Unfair penalization occurs when the test content makes the test more difficult for some students than for others. Bias may occur, for example, if a test includes vocabulary that is unfamiliar to students because of their culture or geographic location. Bias may also occur if the test contains images that are more familiar to one group than another, or demands language skills beyond those of the targeted students. For example, if a reading assessment contains vocabulary related to rural life, then inner city students will potentially be more disadvantaged than rural students. In addition, bias occurs when assessments that are based on letter-sound principles are used with students who do not have access to the sounds of language (i.e., students who are deaf or hard-of-hearing).

Assessment developers typically go to great lengths to make sure assessment items are not biased. Examine the publishers' manual for evidence that item reviews to guard against bias have been conducted.

Validity, reliability and freedom from bias are all necessary conditions for all assessment. They are not interchangeable (Linn and Miller 2005). For example, an assessment may offer consistent results (high reliability) without measuring what was aimed at (low validity); and conversely a measurement with all the hallmarks of validity may not have high reliability. The key points of technical quality are summarized in Figure 8.12.

Figure 8.12. Key Points in Technical Quality of Assessments: Long- and Medium-Cycle Assessments

Technical Quality	Key Points
Validity	<ul style="list-style-type: none"> <li>Assessments need to be valid for the intended purpose</li> <li>The extent to which the information the assessment provides is accurate, adequate, and appropriate for a specific decision-making purpose</li> <li>While people often refer to the "validity of a test," it is more correct to refer to the validity of the <i>interpretations</i> that can be made from the results of a test</li> <li>No test is valid for all purposes</li> </ul>
Reliability	<ul style="list-style-type: none"> <li>Consistency of the test results, repeatedly and over time</li> <li>Results of a test are reliable if they are replicable (despite changes in test administration and scoring, e.g., time of administration or who scores a test)</li> <li>Reliability is important because it is a necessary, but not sufficient condition for validity. If assessment results are not consistent, then it is reasonable to conclude that the scores do not accurately measure what the test is intended to measure</li> </ul>
Freedom from Bias	<ul style="list-style-type: none"> <li>Information or condition in an assessment that unfairly disadvantages a student or groups in showing knowledge in the content</li> <li>An assessment free from bias produces same scores for students at the same attainment level, despite students' demographics (e.g., gender, ethnicity, primary language)</li> <li>Two forms of bias: (1) offensiveness – content offends or upsets particular subgroups, (2) unfair penalization – content more difficult for some students than others</li> </ul>

In the next section, the ideas of validity, reliability and bias are considered in the context of formative assessment practice.

### Technical Quality and Formative Assessment

In formative assessment, the evidence generated by a variety of means is intended to provide information about the students' learning progress in relation to the specific learning goals (i.e., for a lesson) and to be used to inform immediate decisions about next steps in teaching and learning. As alignment to goals is important for annual

and interim assessment, so it is for formative assessment. Teachers will need to be clear about the specific learning goals (what students will learn, *not* what they will do) and what a successful performance entails. For example, learning goals for third grade readers might be to 1) understand that the main idea is the author's message about a topic, minus all the details; and 2) determine the main idea of a text. The performances of understanding and skills for these goals would be for the students to 1) explain the main idea of a text; 2) locate where the author directly expresses the main idea (message) in text; and 3) explain how the important details describe the main idea. The teacher can align her evidence gathering strategies with the goals and performance criteria.

For assessment to be formative it must be both timely and produce information that can inform teaching practice during its ongoing course (Erickson 2007). For this reason the immediate or proximate timing of evidence is a key component of formative assessment validity. In addition, for formative assessment to be valid the resulting information must also yield substantive insights into students' current learning status that can be used in subsequent pedagogical action (Heritage 2013).

An important point about validity in formative assessment concerns the consequences of the assessment use. Because action resulting from the use of formative assessment evidence is intended to produce benefits to student learning, consequences represent an important component of the validity of such assessment. Even if assessments are formative in intention they may not be so in practice if they do not generate further learning (Stobart 2006; Wiliam and Black 1996).

Reliability for classroom formative assessment takes a very different form because errors in instructional decisions can be rectified quickly through gathering more evidence of learning (Shepard 2001). Reliability in relation to instructional decisions can be thought of as "sufficiency of information" (Smith 2003, p. 30). In other words, teachers have to be confident that they have enough information about the student's learning to make a reasonable judgment about the current status of that learning. This idea of sufficiency of information for reliability argues for multiple sources of evidence before a teacher makes an instructional decision. The wider the range of information, and the more frequently the information is collected, the more accurately learning can

be inferred (Griffin, and others 2010). In practical terms, this might mean that before making a judgment about student learning on specific features of language, a teacher has evidence from students' oral language production, from a quick-write, and from a text that has been underlined by the students to identify the specific language feature in question. The more this kind of evidence can be gathered in the context of everyday learning tasks, the less time will be taken away from instruction and the more reliable the evidence gathered about a student's learning will be (Linn and Baker 1996).

Because reading, writing, speaking and listening skills do not develop in lockstep across all students, formative assessment is inevitably personalized and teachers will need to employ strategies that tap into individual's knowledge and skills. Whatever evidence sources a teacher selects, they should account for the range of students present in the class so that all students have the opportunity to show where they are in their learning and have the prospect of moving forward from their current status. For example, well-designed questions and tasks that are sufficiently open-ended can give all students the opportunity to reveal their learning. Similarly, formative assessment should not include any elements that would prevent some students from showing where they are relative to goals.

#### Figure 8.13. Key Points in Technical Quality of Assessments: Short-Cycle Formative Assessments

- Evidence gathered by the teacher is in alignment to specific student learning goals derived from standards
- Evidence gathered needs to be timely and contain information that can inform teaching
- Validity of formative assessment mainly lies in the use of evidence: information gathered must yield substantive insights to students' current learning status that will be used for pedagogical action in order to move students toward achieving learning goals
- Reliability pertains to gathering enough information (e.g., multiple sources) about student learning in order to make a reasonable, accurate judgment for subsequent instructional decisions
- To ensure freedom from bias, evidence gathering should be personalized to students so all students have the opportunity to show where they are in their learning and have the prospect of moving forward from their current learning status.

## Conclusion

The use of assessment by teachers is a critical component of students' achievement of the CA CCSS for ELA/Literacy and the CA ELD Standards. Only when teachers and leaders have a range of accurate information about student learning can they be in a position to make decisions that will advance learning. Key to informing the decisions educators need to make is a balanced and comprehensive system of assessment that provides different levels of detail for different decision-making purposes. Within such an assessment system, districts and school personnel need to strike the right balance in terms of the range of available assessments to teachers from the state or district, to those adopted by individual schools, to assessments embedded in curriculum materials, to ongoing day-by-day formative assessment practices that teachers engage in during instruction. Assessment operates in the service of learning and involves careful consideration of the decisions that teachers need to make, when in the school year they need to make them to ensure student progress, and the assessment tools and processes they need to inform their decision-making. In combination with the right assessments for the right purposes, teachers' skillful use of assessment to support learning are critical to ensure that students in California meet the ambitious language and literacy standards that have been set forth.

## Works Cited

- Abedi, Jamal. 2002. "Assessing and Accommodations of English Language Learners: Issues, Concerns and Recommendations." *Journal of School Improvement* 3 (1): 83–89.
- Abedi, Jamal, and Nancy Ewers. 2013. *Smarter Balanced Assessment Consortium: Accommodations for English Language Learners and Students with Disabilities: A Research-Based Decision Algorithm*.  
<http://www.smarterbalanced.org/wordpress/wp-content/uploads/2012/08/Accommodations-for-under-represented-students.pdf>
- Abedi, Jamal, and Carol Lord. 2001. *Assessment and Accommodations for English Language Learners: Issues and Recommendations*. (CRESST Policy Brief No. 4). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- Allal, Linda. 2010. "Assessment and the Regulation of Learning." In Penelope Peterson, Eva Baker, and Barry McGaw (Eds.), *International Encyclopedia of Education Vol. 3* (3rd ed.), 348–352. Oxford: Elsevier.
- American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME). 1999. *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Ames, Carole, and Jennifer Archer. 1988. "Achievement Goals in the Classroom: Students' Learning Strategies and Motivation Processes." *Journal of Educational Psychology* 80: 260–267.
- Andrade, Heidi G. 2013. *Essay Scoring Rubric*. Unpublished document.
- Andrade, Heidi, Xiaolei Wang, Ying Du, and Robin L. Akawi. 2009. "Rubric-Referenced Self-Assessment and Self-Efficacy for Writing." *The Journal of Educational Research* 102 (4): 287–302.
- Arter, Judith A., and Jan Chappuis. 2006. *Creating and Recognizing Quality Rubrics*. Portland, OR: Educational Testing Service.
- Arter, Judith, and Vicki Spandel. 1992. "Using Portfolios of Student Work in Instruction and Assessment." *Educational Measurement: Issues and Practice* 11 (1): 36–44.

- Bailey, Alison L., and Margaret Heritage. 2008. *Formative Assessment for Literacy, Grades K-6: Building Reading and Academic Language Skills Across the Curriculum*. Thousand Oaks, CA: Corwin/Sage Press.
- Bangert-Drowns, Robert L., Chen-Lin C. Kulik, James A. Kulik, and Mary Teresa Morgan. 1991. "The Instructional Effect of Feedback in Test-like Events." *Review of Educational Research* 61: 213–238.
- Bell, Beverley, and Bronwen Cowie. 2000. "The Characteristics of Formative Assessment in Science Education." *Science Education* 85: 536–553.
- Black, Paul J., and Dylan Wiliam. 1998. "Assessment and Classroom Learning." *Assessment in Education: Principles Policy and Practice* 5: 7–73.
- Black, Paul, Christine Harrison, Clare Lee, Bethan Marshall, and Dylan Wiliam. 2003. *Assessment for Learning: Putting It into Practice*. New York, NY: Open University Press.
- Boekaerts, Monique. 2006. Self-Regulation and Effort Investment. In K. A. Renninger and I.E. Siegel (Eds.). *Handbook of Child Psychology: Vol. 4. Child Psychology in Practice* (6th ed.), 345–377. New York: Wiley.
- Bowles, Melissa, and Charles W. Stansfield. 2008. *A Practical Guide to Standards-Based Assessment in the Native Language* (NLA—LEP Partnership). [http://www.ncela.us/files/rcd/BE024208/A\\_Practical\\_Guide.pdf](http://www.ncela.us/files/rcd/BE024208/A_Practical_Guide.pdf) (accessed June 10, 2013).
- Brookhart, Susan M. 2013. *How to Create and Use Rubrics for Formative Assessment and Grading*. Alexandria, VA: ASCD.
- Bunch, George C., Amanda Kibler, and Susan. Pimental. 2012. "Realizing Opportunities for English Learners in the Common Core English Language Arts and Disciplinary Literacy Standards." In Kenji Hakuta and M. Santos (Eds.), *Understanding Language: Commissioned Papers on Language and Literacy Issues in the Common Core State Standards and Next Generation Science Standards*, 1–16. Palo Alto, CA: Stanford University.
- Butler, Deborah L., and Philip H. Winne. 1995. "Feedback and Self-Regulated Learning: A Theoretical Synthesis." *Review of Educational Research* 65 (3): 245–281.

- California Department of Education. 2013. *California Common Core State Standards: English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects*. Sacramento: California Department of Education.  
<http://www.cde.ca.gov/be/st/ss/documents/finalelaccsstandards.pdf> (accessed April 5, 2013).
- California Department of Education. 2014 *California English Language Development Standards*. Sacramento: California Department of Education.
- Carnegie Council on Advancing Adolescent Literacy. 2010. *Time to Act: An Agenda for Advancing Adolescent Literacy for College and Career Success*. New York: Carnegie Corporation of New York.
- Chappuis, Jan, Rick J. Stiggins, Steve Chappuis, and Judith. A. Arter. 2012. *Classroom Assessment for Student Learning: Doing It Right – Using It Well* (2nd ed.). Portland, OR: Pearson Assessment Training Institute.
- Christie, Frances. (Ed.). 1999. *Pedagogy and the Shaping of Consciousness: Linguistic and Social Processes*. London, UK: Cassell Academic.
- Cohen, Nicole, Tracey E. Hall, Ge Vue, and Patti Ganley. 2011. “Becoming Strategic Readers: Three Cases Using Formative Assessment, UDL, and Technology to Support Struggling Middle School Readers.” In Pendred E. Noyce and Daniel T. Hickey (Eds.), *New Frontiers in Formative Assessment*, 129–140. Cambridge, MA: Harvard Education Press.
- Council of Chief State School Officers (CCSSO) Formative Assessment State Collaborative. 2006. *Definition of Formative Assessment*. Washington, DC: CCSSO.
- Crooks, Terence J. 1988. “The Impact of Classroom Evaluation on Students.” *Review of Educational Research* 58: 438–481.
- Derewianka, Beverly. 2011. *A New Grammar Companion for Teachers*. Sydney, NSW: Primary English Teaching Association.
- Dweck, Carol S. 1999. *Self-Theories: Their Role in Motivation, Personality and Development*. Philadelphia, PA: Psychology Press.
- Dweck, Carol S. 2006. *Mindset: The New Psychology of Success*. New York: Random House.

- Erickson, Frederick. 2007. "Some Thoughts on 'Proximal' Formative Assessment of Student Learning." *Yearbook of the National Society for the Study of Education* 106: 186–216.
- Fang, Zhihui and Zhujun Wang. 2011. "Beyond Rubrics: Using Functional Language Analysis to Evaluate Student Writing." *Australian Journal of Language and Literacy* 34 (2): 147–165.
- Gibbons, Pauline. 2009. *English Learners, Academic Literacy, and Thinking: Learning in the Challenge Zone*. Portsmouth, NH: Heinemann.
- Griffin, Patrick. 2007. The Comfort of Competence and the Uncertainty of Assessment. *Studies in Educational Evaluation* 33: 87–99.
- Griffin, Patrick, Leanne Murray, Esther Care, Amanda Thomas, and Pierina Perri. 2010. "Developmental Assessment: Lifting Literacy Through Professional Learning Teams." *Assessment in Education: Principles, Policy and Practice* 17 (4): 383–397.
- Hacker, Douglas J., John Dunlosky, and Arthur C. Graesser (Eds). 1998. *Metacognition in Educational Theory and Practice*. Mahwah, NJ: Lawrence Erlbaum.
- Harlen, Wynne, and Ruth Deakin Crick. 2002. *A Systematic Review of the Impact of Summative Assessment and Tests on Student Motivation for Learning*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.  
<http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=Pbyl1CdsDJU%3D&tabid=106&mid=1967> (accessed February 2013).
- Harlen, Wynne, and Mary James. 1997. "Assessment and Learning: Differences and Relationships between Formative and Summative Assessment." *Assessment in Education: Principles, Policy and Practice* 4 (3): 365–379.
- Harlen, Wynne. 2007. "Formative Classroom Assessment in Science and Mathematics." In J. H. McMillan (Ed.), *Formative Classroom Assessment: Theory into Practice*, 116–135. New York: Teachers College Press.
- Hattie, John, and Helen Timperley. 2007. "The Power of Feedback." *Review of Educational Research* 77 (1): 81–112.
- Heritage, Margaret. 2010. *Formative Assessment: Making It Happen in the Classroom*. Thousand Oaks, CA: Corwin Press.

- Heritage, Margaret. 2013. *Formative Assessment in Practice: A Process of Inquiry and Action*. Cambridge, MA: Harvard Education Press.
- Heritage, Margaret. 2014. Formative Assessment as the Key to Instructional Practice. Webinar provided by the Regional Educational Laboratory, Mid-Atlantic. June 19, 2014.
- Herman, Joan L. 2010. *Coherence: Key to Next Generation Assessment Success*. Los Angeles, CA: CRESST.  
[http://www.cse.ucla.edu/products/policy/coherence\\_v6.pdf](http://www.cse.ucla.edu/products/policy/coherence_v6.pdf)
- Herman, Joan L., Pamela R. Aschbacher, and Lynn Winters. 1992. *A Practical Guide to Alternative Assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Herman, Joan L., and Margaret Heritage. 2007. *Moving from Piecemeal to Effective Formative Assessment Practice: Moving Pictures on the Road to Student Learning*. Paper presented at the Council of Chief State School Officers Assessment Conference, Nashville, TN.
- Herman, Joan L., Margaret Heritage, and Pete Goldschmidt. 2011. *Developing and Selecting Assessments of Student Growth for Use in Teacher Evaluation Systems*. Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Herman, Joan L., Ellen Osmundson, and Ronald Dietel. 2010. *Benchmark Assessment for Improved Learning* (AACC Report). Los Angeles, CA: University of California.
- International Reading Association. 2000. *Teaching All Children to Read: The Roles of the Reading Specialist: A Position Statement of the International Reading Association*. [http://www.reading.org/Libraries/position-statements-and-resolutions/ps1040\\_specialist.pdf](http://www.reading.org/Libraries/position-statements-and-resolutions/ps1040_specialist.pdf)
- Jenkins, Joseph R. 2003. *Candidate Measures for Screening At-Risk Students*. Paper presented at the National Research Center on Learning Disabilities Responsiveness-to-Intervention symposium, Kansas City, MO.

- Joseph, Laurice M. 2002. "Best Practices in Planning Intervention for Students with Reading Problems." In Alex Thomas, and Jeff Grimes (Eds.), *Best Practices in School Psychology IV*, 803–816. Bethesda, MD: National Association of School Psychologists.
- Kluger, Avraham N., and Angelo DeNisi. 1996. "The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory." *Psychological Bulletin* 119: 254–284.
- Lane, Suzanne. 2013. "Performance Assessment." In James H. McMillan. (Ed). *Sage Handbook of Research on Classroom Assessment*, 313–331. Thousand Oaks, CA: SAGE Publications.
- Linn, Robert L., and Eva L. Baker. 1996. "Can Performance-Based Student Assessment Be Psychometrically Sound?" In Kenneth J. Rehage, Joan Boykoff Baron, and Dennie Palmer Wolf (Eds.), *Performance-Based Assessment: Challenges and Possibilities: Yearbook of the National Society for the Study of Education* 95 (1): 84-103). Chicago, IL: University of Chicago Press for the National Society for the Study of Education.
- Linn, Robert L., and Michael David Miller. 2005. *Measurement and Assessment in Teaching*. Upper Saddle River, NJ: Pearson Education.
- Linquanti, Robert. 2014. *Supporting Formative Assessment for Deeper Learning: What Policymakers Must Know and Be Able to Do*. Paper prepared for the Formative Assessment for Student and Teachers (FAST) State Collaborative on Assessment and Student Standards (SCASS). Washington, DC: Council of Chief State School Officers (CCSSO).
- Lodico, Marguerite G., Elizabeth S. Ghatala, Joel R. Levin, Michael Pressley, and John A. Bell. 1983. "The Effects of Strategy-Monitoring Training on Children's Selection of Effective Memory Strategies." *Journal of Experimental Child Psychology* 35 (2): 263–277.

McManus, Sarah. 2008. *Attributes of Effective Formative Assessment*. Paper prepared for the Formative Assessment for Student and Teachers (FAST) and State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers (CCSSO).

[http://www.ccsso.org/Resources/Publications/Attributes\\_of\\_Effective\\_Formative\\_Assessment.html](http://www.ccsso.org/Resources/Publications/Attributes_of_Effective_Formative_Assessment.html).

National Association of State Directors of Special Education (NASDSE). 2005.

*Response to Intervention: Policy Considerations and Implementation*. Alexandria, VA: Author.

National Research Council. 1999. *Improving Student Learning: A Strategic Plan for Education Research and its Utilization*. Washington, DC: National Academy Press.

National Research Council. 2001. *Knowing What Students Know: The Science of Design and Educational Assessment*. Washington, DC: National Academy Press.

National Research Council. 2012. *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. Washington, DC: National Academy Press.

National Research Center on Learning Disabilities (NRCLD). 2006. *Progress Monitoring*. [http://www.nrclld.org/rti\\_manual/pages/RTIManualSection2.pdf](http://www.nrclld.org/rti_manual/pages/RTIManualSection2.pdf)

Organisation for Economic Co-operation and Development (OECD). 2005. *Formative Assessment: Improving Student Learning in Secondary Classrooms*. Centre for Educational Research and Innovation.

<http://www.oecd.org/dataoecd/19/31/35661078.pdf> (accessed May 9, 2008).

Palm, Torulf. 2008. "Performance Assessment and Authentic Assessment: A Conceptual Analysis of the Literature." *Practical Assessment, Research, and Evaluation* 13 (4): 1–11.

Pennock-Roman, Maria, and Charlene Rivera. 2011. "Mean Effects of Test Accommodations for ELLs and Non-ELLs: A Meta-Analysis of Experimental Studies." *Educational Measurement: Issues and Practice* 30 (3), 10–28.

- Perie, Mariane, Scott Marion, and Brian Gong. 2009. "Moving Toward a Comprehensive Assessment System: A Framework for Considering Interim Assessments." *Educational Measurement: Issues and Practice* 28 (3): 5–13.
- Perry, Nancy, Karen O. VandeKamp, Louise K. Mercer, and Carla J. Nordby. 2002. "Investigating Teacher-Student Interactions that Foster Self-Regulated Learning." *Educational Psychologist* 37 (1): 5–15.
- Pintrich, Paul R. 2000. "The Role of Goal Orientation in Self-Regulated Learning." In Monique Boekaerts, Paul R. Pintrich, and Moshe Zeidner (Eds.), *Handbook of Self-Regulation*, 451–502. San Diego, CA: Academic Press.
- Popham, W. James. 1995. *Classroom Assessment: What Teachers Need to Know*. Boston: Allyn and Bacon.
- Popham, W. James. 2010. *Transformative Assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Poppers, Audrey E. 2011. "Identifying Craft Moves: Close Observation of Elementary Students' Writing." In Pendred E. Noyce and Daniel T. Hickey (Eds.), *New Frontiers in Formative Assessment*, 89–107. Cambridge, MA: Harvard Education Press.
- Puckett, Margaret B., and Deborah Diffily. 2004. *Teaching Young Children: An Introduction to the Early Childhood Profession* (2nd ed.). Clifton Park, NY: Delmar Learning.
- Resnick, Lauren B., and Daniel P. Resnick. 1992. "Assessing the Thinking Curriculum: New Tools for Educational Reform." In Bernard R. Gifford and Mary Catherine O'Connor (Eds.), *Changing Assessments: Alternative Views of Aptitude, Achievement and Instruction*, 37–76. New York: Kluwer/Springer Science+Business Media.
- Ruiz-Primo, Maria A., and Erin M. Furtak. 2004. *Informal Assessment of Students' Understanding of Scientific Inquiry*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Ruiz-Primo, Maria A., and Erin M. Furtak. 2006. "Informal Formative Assessment and Scientific Inquiry: Exploring Teachers' Practices and Student Learning." *Educational Assessment* 11 (3 and 4): 237–263.

- Ruiz-Primo, Maria A., and Erin M. Furtak. 2007. "Exploring Teachers' Informal Formative Assessment Practices and Students' Understanding in the Context of Scientific Inquiry." *Journal of Educational Research in Science Teaching* 44 (1): 57–84.
- Sadler, D. Royce. 1989. Formative Assessment and the Design of Instructional Strategies. *Instructional Science* 18: 119–144.
- Schunk, Dale H., and Barry J. Zimmerman. 2008. *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. Mahwah, NJ: Lawrence Erlbaum.
- Smarter Balanced Accessibility Consortium (SBAC). 2013a. <http://www.smarterbalanced.org/>
- Smarter Balanced Accessibility Consortium (SBAC). 2013b. *Smarter Balanced Accessibility and Accommodations: Meeting the Needs of All Students*. [http://www.smarterbalanced.org/wordpress/wp-content/uploads/2012/07/SmarterBalanced\\_Accessibility\\_Factsheet.pdf](http://www.smarterbalanced.org/wordpress/wp-content/uploads/2012/07/SmarterBalanced_Accessibility_Factsheet.pdf)
- Smarter Balanced Accessibility Consortium (SBAC). 2014. *Smarter Balanced Assessment Consortium: Usability, Accessibility, and Accommodations Guidelines*. [http://www.smarterbalanced.org/wordpress/wp-content/uploads/2013/09/SmarterBalanced\\_Guidelines\\_091113.pdf](http://www.smarterbalanced.org/wordpress/wp-content/uploads/2013/09/SmarterBalanced_Guidelines_091113.pdf)
- Shepard, Lorrie A. 2000. "The Role of Assessment in a Learning Culture." *Educational Researcher* 29 (7): 4–14.
- Shepard, Lorrie A. 2001. The Role of Classroom Assessment in Teaching and Learning. In Virginia Richardson (Ed.), *Handbook of Research on Teaching* (4th ed.), 1066–1101. Washington, DC: American Educational Research Association.
- Shepard, Lorrie A. 2005. "Linking Formative Assessment to Scaffolding." *Educational Leadership* 63 (3): 66–71.
- Shepard, Lorrie A. 2005. *Will Commercialization Enable or Destroy Formative Assessment?* (Original title: Formative assessment: Caveat emptor). Paper presented at the ETS Invitational Conference 2005: The Future of Assessment, Shaping Teaching and Learning, New York.
- Smith, Jeffrey K. 2003. "Reconsidering Reliability in Classroom Assessment and Grading." *Educational Measurement: Issues and Practices* 22 (4): 26–34.

- Soltero-Gonzalez, Lucinda, Kathy Escamilla, and Susan Hopewell. 2012. "Changing Teachers' Perceptions About the Writing Abilities of Emerging Bilingual Students: Towards a Holistic Bilingual Perspective on Writing Assessment." *International Journal of Bilingual Education and Bilingualism* 15 (1): 71–94.
- Spycher, Pamela, and Karin Linn-Nieves. 2014. Reconstructing, Deconstructing, and Constructing Complex Texts. In Pamela Spycher (Ed.) *The Common Core State Standards in English Language Arts/Literacy for English Language Learners: Grades K–5*. Alexandria, Virginia: TESOL Press.
- Stansfield, Charles W., and Melissa Bowles. 2006. "Study 2: Test Translation and State Assessment Policies for English Language Learners." In Charlene Rivera and Eric Collum (Eds.), *State Assessment Policy and Practice for English Language Learners: A National Perspective*, 177–313. Mahwah, NJ: Lawrence Erlbaum/Routledge.
- Stobart, Gordon. 2006. "The Validity of Formative Assessment." In John Gardner (Ed.), *Assessment and Learning*, 133–146. London, UK: Sage.
- Understanding Language. 2013. *Instructional Unit: Persuasion Across Time and Space*. Palo Alto, CA: Stanford University.  
[http://ell.stanford.edu/sites/default/files/ela\\_archives/understanding\\_language\\_materials\\_Jan2013.pdf](http://ell.stanford.edu/sites/default/files/ela_archives/understanding_language_materials_Jan2013.pdf)
- Venn, John J. 2000. *Assessing Students with Special Needs* (2nd ed.). Upper Saddle River, NJ: Merrill.
- Wiliam, Dylan, and Paul J. Black. 1996. "Meanings and Consequences: A Basis for Distinguishing Formative and Summative Functions of Assessment?" *British Educational Research Journal* 22 (5): 537–548.
- Wiliam, Dylan. 2006. *Does Assessment Hinder Learning?* Paper presented at the ETS Invitational Seminar at the Institute of Civil Engineers, London, UK.
- Wiliam, Dylan. 2007. "Keeping Learning on Track: Classroom Assessment and the Regulation of Learning." In Frank K. Lester Jr. (Ed.), *Second Handbook of Mathematics Teaching and Learning*, 1053–1098. Greenwich, CT: Information Age Publishing.

Wiliam, Dylan. 2011. "What is Assessment for Learning?" *Studies in Educational Evaluation* 37: 3–14.